```
Nephila clavipes s
Spider silk protei
Spider natural sil
Nephila clavipes s
N. clavipes spider
Spider recombinant
N. clavipes spidro
N. clavipes spidro
N. clavipes spidro
Spider dragline va
N. clavipes spidro
Spider dragline va
Polymer of an anal
Dragline protein i
                                                                                                                                                                                                                   N. clavipes spidro
N. clavipes spidro
Spider silk protei
PMISSI MiSP spider
Epstein Barr Virus
Epstein Barr Virus
EBV tethering prot
New DNA sequence I
                                                                                                                                                                                                                                                                                                         FLGA Gly-ala inser
N. clavipes spidro
N. clavipes spidro
Glycine-rich repea
Epstein Barr virus
Spider allk protei
N.clavipes dragilin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Spider; orb-weaver; silk protein; pQE(spl)7; structural protein;
purification; fibre; spinning.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Spider recombinant silk protein pQE(spl)7
                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note- "encoded by GGY'
                                                                                                                                                                                    AAU11798
AAM50044
                                                                                                                                                                                                                                                                                                                   AAM50048
AAM50040
AAW05704
AAW79126
AAY40098
                                                                                                                     AAR99055
AAY40101
AAY40102
AAU11793
AAU11797
                                                                                                                                                                                                                                                                 AAY28843
AAY95856
                                                                                                                                                                                                                                            AAY40099
                                                                                                                                                                                                                                                     AAR80168
                                                                                                           AAM50046
                                                                                                                                                                                                           ABB76672
                                                                                                                                                                                                                      AAM50035
                                                                                                                                                                                                                                AAM50041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Butler MM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB82608 standard; Protein; 291 AA
  01-NOV-2000; 2000WO-US30086.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20-JAN-2000; 2000US-0490291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Arcidiacono S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (MELL/) MELLO C M.
(ARCI/) ARCIDIACONO S.
(BUTL/) BUTLER M M.
(USSA ) US SEC OF ARMY.
 Misc-difference 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nephila clavipes.
             441.1
339.0
339.1
337.8
37.8
37.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO20015333-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26-JUL-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mello CM,
1031
1015
1015
1015
1015
10018
10008
974
977
976
976
976
                                                                                                                                                                                                         859.5
69.5
69.5
677.5
610.5
610.5
610.5
593.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB82608;
                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB82608
 Spider recombinant
Spider recombinant
Spider recombinant
N. clavipes spidro
Spider dragline va
Polymer of an anal
N. clavipes spidro
N. clavipes spidro
N. clavipes spidro
                                                                                    (without alignments)
1995.696 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                          1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseqy-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseqy-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseqy-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseqy-embl/AA1985.DAT:*
6: /SIDS2/gcgdata/geneseqy-embl/AA1986.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1980.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1980.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1980.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1990.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1990.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1992.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1993.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1993.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1993.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1993.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1993.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1995.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1995.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1995.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1995.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1995.DAT:*
7: /SIDS2/gcgdata/geneseqy-embl/AA1995.DAT:*
                                                                        December 18, 2002, 16:05:10 ; Search time 19.4298 Seconds
                                                                                                                             1485
1 MRGSHHHHHHGSMASGRGGL......YGGLGSQGTSGIRRPAAKLN 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /SIDSZ/gcgdata/geneseg/genesegp-embl/AA2000.DAT:*/SIDSZ/gcgdata/geneseg/genesegp-embl/AA2001.DAT:*/SIDSZ/gcgdata/geneseg/genesegp-embl/AA2002.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /SIDS2/gcgdata/geneseg/genesegp-embl/AA1997.DAT:*/SIDS2/gcgdata/geneseg/genesegp-embl/AA1999.DAT:*/SIDS2/gcgdata/geneseg/genesegp-embl/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
          GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd
                                                                                                                                                                                                                 hits satisfying chosen parameters:
                                                                                                                                                                                             908470 seqs, 133250620 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES

    protein search, using sw model

                                                                                                                                                                                                                                                                                             Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AABB2609
AABB2609
AABB2610
AAM50038
AAR99053
AAY40100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAM50042
AAM50047
AAM50037
AAM50039
                                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                   100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                  A_Geneseq_101002:*
                                                                                                                                                                                                                                        seq length: 0
seq length: 2000000000
                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length DB
                                                                                                                   US-09-490-291-2
                                                                                                                                                                                                                                                                                  Maximum Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          291
681
691
989
606
606
630
676
1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0
96.0
91.4
69.8
69.8
69.8
69.5
69.5
69.5
```

1485 1425 1357 1037 1036 1036 1031.5 1031.5 1031.5

22 4 3 2 5 4 5 1 0 1 0 1 0 1

Score

Result

Total number of

Searched:

Minimum DB Maximum DB

Database

Title: Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

```
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                           The present sequence is that of orb-weaver spider (Nephila clavipes)
recombinant silk protein, pQE(spl)? The invention provides
methods for purifying and spinning spider silks and other
structural proteins. Organic acids are used to lyse recombinant
cells or other biological samples (such as non-recombinantly
derived cells), and enrich the purity and yields of structural
cells or other biological samples (such as non-recombinantly
the structural proteins intect. In the case of silk proteins, the
proteins by hydrolysing many of the macromolecules while leaving
the structural proteins intect. In the case of silk proteins, the
cresulting lysate is further purified by non-exchange or affinity
chromatography and processed into an aqueous-based mixture for
fibre spinning. In the present case, the pQE(spl) gene was clone
fibre spinning. In the present case, the pQE(spl) gene was clone
fibre spinning. In the present case, the pQE(spl)
and recombinant silk protein was purified using formic acid and
into vector pQE's for recombinant expression in Escherichia coli,
and recombinant silk protein was purified using formic acid and
including films, fibres, woven articles, sutures, ballistic
cof the invention can be used in the construction of many materials
including films, fibres, woven articles, sutures, ballistic
cof the invention and smaller volumes of reagents, results in
better recovery of protein at higher purity (70-99%), is easy to
scale up, and the fibres are spun in an environmentally benign
                                                  Recovering structural polypeptides in a biological sample, useful for purifying and spinning spider silks and other structural proteins, comprises treating the sample containing the polypeptides with an acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 QGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MRGSHHHHHHGSMASGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGIRRPAAKLN 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Spider; orb-weaver; silk protein; pQE((SP1)4/(SP2)1)4; structural protein; purification; fibre; spinning.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Spider recombinant silk protein pQE((SP1)4/(SP2)1)4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100.0%; Score 1485; DB 22;
100.0%; Pred. No. 5.3e-108;
tive 0; Mismatches 0;
                                                                                                                       Claim 2; Page 33-34; 49pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB82609 standard; Protein; 681 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 291; Conservative
            WPI; 2001-483136/52.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                               291 AA;
                          N-PSDB; AAH26301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              02-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB82609;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB82609
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
```

```
The present sequence is that on the weaver spinler (Weapila Clavipes) recombinant silk protein pGE(SP1)4(SP2)1)4. The invention provides methods for purifying and spinning spider silks and other structural proteins. Organic acids are used to lyse recombinant cells or other biological samples (such as non-recombinantly derived cells), and enrich the purity and yields of structural proteins many of the macromolecules while leaving the structural proteins intact. In the case of silk proteins, the resulting lysate is further purified by ion-exchange or affinity chromography and processed into an aqueous-based mixture for fibre spinning. In the present case, the pQE((SP1)4/(SP2)1)4 gene was cloned into vector pQE-9 for recombinant expression in the propionic acid and anion-exchange chromatography on QAE-Sephadex A50. Products obtained using the methods of the including films, fibres, woven articles, sutures, ballistic products and parachute cords. The new method has the fibre load than an articles, sutures, ballistic required load than an articles of the products of the products of the products of the following advantages over prior articles, the policy of the products of the products of the following advantages over prior articles, the policy of 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Recovering structural polypeptides in a biological sample, useful for purifying and spinning spider silks and other structural proteins, comprises treating the sample containing the polypeptides with an acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 12; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                present sequence is that of orb-weaver spider (Nephila clavipes)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MRGSHHHHHHGSMASGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      requires less time and smaller volumes of reagents, results in better recovery of protein at higher purity (70-9948), is easy tscale up, and the fibres are spun in an environmentally benign solution reducing hazardous waste accumulation and cost.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 681;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 GGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 1425; DB 22;
Pred. No. 4.7e-103;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Butler MM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 2; Page 34-37; 49pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96.0%;
                                                                                                                                                                                                                                                                                                                                                          20-JAN-2000; 2000US-0490291.
                                                                                                                                                                                                                                                                  31-NOV-2000; 2000WO-US30086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mello CM, Arcidiacono S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 96.03
Best Local Similarity 95.93
Matches 282; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US SEC OF ARMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MELL/) MELLO C M. (ARCI/) ARCIDIACONO S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-483136/52.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BUTLER M M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      681 AA;
Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAH26302
                                                                                   WO200153333-A1.
                                                                                                                                                                          26-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (ARCI/)
(BUTL/)
(USSA )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         169
```

```
The present sequence is that of orb-weaver spider (Nephila clavipes)

recombinant silk protein pET((SP1)4/(SP2)1)4. The invention

provides methods for purifying and spinning spider silks and other

crels or other biological samples (such as non-recombinantly

cells or other biological samples (such as non-recombinantly

cells or other biological samples (such as non-recombinantly

cells or other biological samples (such as non-recombinantly

the structural proteins intact. In the case of silk proteins, the

resulting lygate is further purified by lon-exchange or affinity

chromatography and processed into an aqueous-based mixture for

thromatography and processed into an aqueous-based mixture for

fibre spinning. In the present case, the pET((SP1)4/(SP2)1)4 gene

was cloned into vector pET34 for recombinant was obtained in

Escherichia coll, and recombinant silk protein was obtained

Scherichia coll, when denaturant (3 M quanidine-HCI) was added

to the propionic acid, the recombinant protein was obtained

to the propionic acid, the recombinant protein was obtained

purity. Products obtained using the methods of the invention can

cut propion articles, stutures, ballistic protection, parachutes

and parachute conds. The new method has the following advantages

cover prior art: it involves fewer steps, requires less time and

smaller volumes of reagents, results in better recovery of protein

at higher purity (70-99%), is easy to scale up, and the fibres are

source miniationmentally benign solution reducing hazardous waste
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Recovering structural polypeptides in a biological sample, useful for purifying and spinning spider silks and other structural proteins, comprises treating the sample containing the polypeptides with an acid
                    AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSG
                                                                                                                                                                                                                                                                                                 Spider; orb-weaver; silk protein; pET((SP1)4/(SP2)1)4; structural protein; purification; fibre; spinning.
                                                                                                                                                                                                                                                            Spider recombinant silk protein pET((SP1)4/(SP2)1)4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Butler MM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 2; Page 38-40; 49pp; English.
                                                                                                                                      AAB82610 standard; Protein; 691 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-2000; 2000WO-US30086.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20-JAN-2000; 2000US-0490291.
                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Arcidiacono S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (MELL/) MELLO C M.
(ARCI/) ARCIDIACONO S.
(BUTL/) BUTLER M M.
(USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           accumulation and cost.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2001-483136/52.
                                                                                                                                                                                                                                                                                                                                                           Nephila clavipes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAH26303.
                                                                                                                                                                                                                                                                                                                                                                                                   WO200153333-A1.
                                                                                                                                                                                                                     02-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                          26-JUL-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mello CM,
                                                                                                                                                                             AAB82610;
                                                                                                                   AAB82610
ò
```

691 AA;

Sequence

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     spider silk protein, comprising modules, each comprising a group of a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila claupes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (1) as adhesion surfaces for culturing cells; and (iii) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New DNA encoding synthetic spider slik protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOISM12.
                                                                                                                                                                           71 AAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSG 130
                                                                                                                         RGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQG------TSGRGGLGGQGAG 178
                                                                                                                                                                                                                  Gaps
                                                                    This invention describes a novel DNA sequence, encoding a synthetic
                             12;
  Length 691;
                             Indels

    N. clavipes spidroin synthetic homologue SOISM12 protein.

                                                                                                                                                                                                                                                                     ö
 DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
 Score 1357; DB 22
Pred. No. 8.8e-98;
                          0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Guehrs
                                                                                                                                                                                                                                                                                                                                                                 AAM50038 standard; Protein; 989 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 22; Page 34-38; 88pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Scheller J, Conrad U, Grosse F,
91.4%;
95.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09-JUN-2000; 2000DE-1028212.
24-OCT-2000; 2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21-MAR-2001; 2001DE-1013781
                                                                                                                                                                                                                                                                                                                                                                                                                   18-SEP-2002 (first entry)
                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2002-123561/17.
N-PSDB; ABL61039.
               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DE10113781-A1
             Best Local Simi
Matches 272;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-DEC-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic.
 Query Match
                                                                                                                                                                                                                                                                                                                                                                                           AAM50038;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     proteins
                                                                             13
                                                                                                                                                             131
                                                                                                                                                                                                                                           193
                                                                                                                                                                                                                  179
                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                        ò
                                                                                                                                 a
                                                                                                                                                             à
                                                                                                                                                                                      g
                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                              요
```

4

94WO-US06689

```
15-JUN-1994;
     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                 13;
                                          to simplify
Also they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and purified. The modular construction of the invention facilitates incorporation of additional peptide-encoding sequences, e.g. to simplification or modulate solubility. This sequence represents the synthetic N. clavipes spidroin-1 homologue SOISMI2 described in the
                                                                                                                                                                                                  702
                                                                                                                                                            164 QGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGL 523
                                                                                                                                                                                                                                                                    --AAAAAAGGAGGGYGGLGSQGT----- 128
                                                                                                                                                                                                                                                                                                                                                                                         -----AAAAAGGAGQG 232
                                                                                                                                                                                                                                                                                                                                                                                                                                           GYGGLGSQG-AGRGGLGGQGAGAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAGG 820
                                                                                                                                                                                                                                                                                                                                                GGLGSQGTSGRGGLGGQGAGAAA---------AAAAAAGGAGQGGYGG 198
                                                                                                                                                                                                                                                                                                                                                                                                         703 IGSQG-AGRGGIGGQGAGAAAAAAGGAGQGGYGGIGSQGAGRGGQGAGAAAAAAGGAGQG 761
                                                                                                                                 Gaps
                                                                                                                                                  44
                                                                                                                                                                                        ------AAAAAGGAG 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Spider; dragline protein; variant; monomer; polymer; fibre forming region; Spidroin 1; Nephila clavipes; DP1; mimic; DP-1A analogue; fibre; high tensile strength; elasticity; clothing; rope; surgical suture; implant; reinforcement; film; coating.
                                                                                                                                                                                                                               TSSECT SCREET
                                                                                                                                                   -----AAAAAAGGAGOGGY
                                                                                                                                                                                                                                                                               584 AAAAAAGGAGGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGGGGYGGLGGQGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                   Indels 228;
                                                                                                          Length 989;
                                                                                                           23;
                                                                                                                             ..
                                                                                                          Score 1037; DB 2
Pred. No. 7e-73;
7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                            GYGGLGSQGTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Spider dragline variant, DP-1A.9 polymer.
                                                                                                                                                                                       GGLGSQGTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR99053 standard; Protein; 606 AA.
                                                                                                                                                                                                                                                                                                                                                                                       LGSQGTSGRGGLGGOGAGAAA-----
                                                                                                                                                                                                                            OGGYGGLGSQGT-----
                                                                                                                                                11 GSMASGRGGLGGQGAGAAA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                             Matches 250; Conservative
                                                                                                                                                                                                                                                                   GGQGAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  267 AGGGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                       989 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17-JAN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO9429450-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22-DEC-1994.
                                                                     Invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic.
                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR99053;
                                                                                                          Query Match
                                                                                                                      Best Local
                                                                                                                                                                                      45
                                                                                                                                                                                                                            79
                                                                                                                                                                                                                                                                   97
                                                                                                                                                                                                                                                                                                                                                                                      199
                                                                                                                                                                                                                                                                                                          129
                                                                                                                                                                                                                                                                                                                                                                                                                            233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR99053
8888888888888
                                                                                                                                                 à
                                                                                                                                                                  셤
                                                                                                                                                                                     à
                                                                                                                                                                                                         g
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                     윱
                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
```

```
This sequence represents a synthetic spider dragline variant polymer, CC DP-1A.9. The sequence of the DP-1A.9 monomer is given in AAR99052. CT Pe polypeptide monomer is a variant based on a consensus sequence derived from the fibre forming regions of spider dragline protein, csp. the natural protein 1 (Spidroin 1) from Nephila clavipes.

CC derived from the fibre forming regions of spider dragline protein, csp. the natural protein 1 (Spidroin 1) from Nephila clavipes.

CC DNA sequence encoding the monomer may be used in the recombinant production of the variant protein in a recombinant host, e.g. E. coli or Bacillus subtilis. Synthetic analogues of DP1 were designed to mimic the repeating consensus sequence of the natural protein and the pattern of variation among individual repeats. DP-1A analogues are composed of a tandemly repeated 101 amino acid monomer which comprises four crepeats which differ from the consensus sequence given in AAW06201, according to the pattern (1)-(5) given below. This 101 amino acid monomer is repeated 1-16 times in a series of analogue proteins. The individual repeated 1-16 times in a series of analogue proteins.

(1) the poly-alanine sequence encompassing AGRGGGGGGGARGANGG; colloci state surrounding sequence encompassing AGRGGGGGGGARGANGG; colloci subterior of GTG is a repeat multiples of three consecutive residues; (1) aside from the poly-alanine sequence; deletions usually compass integral multiples of three consecutive residues; (2) a repeat in which the entire poly-alanine sequence is confidence is generally preceded by a repeat containing six alanine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The proteins may be used to produce fibres of high tensile strength and elasticity, suitable for clothing, rope, surgical sutures, blomaterials for implants, plastic reinforcements, films, coatings, etc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             50 QGT--SGRGGLGGQGA---GAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             349 AAAAAGGAGGGLGSQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGR 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --- NAAAAAGGAGGG 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --- AAAAAAGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           230 GSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----AAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGA---GAAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---AAAAAAGGAGQGGYGGLGSQGT--SGRGGLGGQGA--
                                                                                                                                                                                                             New synthetic variants of spider dragline protein - for making fibres useful as clothing, surgical silk, plastic reinforcement etc., also related DNA, vectors and transformed cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 1036; DB 16;
Pred. No. 5.6e-73;
4; Mismatches 23;
                                                                                                                                                                                                                                                                                                                                                            Disclosure; Page 85-87; 168pp; English.
       ij.
       ы
       ဥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GYGGLGSQGTSGRGGLGGQGAGAAA-
(DUPO ) DU PONT DE NEMOURS &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             69.8%;
64.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11 GSMASGRGGLGGQGAGAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 64.9
Matches 239; Conservative
                                                                                                                                           WPI; 1995-036479/05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           105 AA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               606 AA;
                                                                     Fahnestock SR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          233
```

ហ

```
Claim 22; Page 48-50; 88pp; German.
                                                                                                                                                                                                                                      AAM50042 standard; Protein; 630 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Grosse F,
                                                                                                                                                                                                                                                                                                                                                                                                               21-MAR-2001; 2001DE-1013781.
                                                                                                                                                                                                                                                                                                                                                                                                                                  09-JUN-2000; 2000DE-1028212.
24-OCT-2000; 2000DE-1053478.
                                                                                                                                                                                                                                                                                  18-SEP-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2002-123561/17.
                                                                                                                                                                            587 YGGLGSQG 594
                                                                                                                                                              272 YGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; ABL61043.
                                                                                                                                                                                                                                                                                                                                                                    DE10113781-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                        13-DEC-2001
                                                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                                             AAM50042;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  proteins
                                  140
                                                                                                                                                                                                                                                             g
                                                                                                                                        g
                                                                                                                                                              ò
                                                                          ò
                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The present sequence represents a polymer of an analogue of the spider silk protein spidrolne major 1. The protein improves the moisturizing/softening action of the compositions. The protein, and its fragments are used in cosmetic or dermatological compositions. These compositions have use as hair or skin care products; and make-up or sunscreens. As the protein is a good, persistent film-formers on the skin of low surface density, it can be used for delivery of active agents that are generally difficult to administer, e.g. vitamins, hormones, moisturizers or agents for treating disorders of the skin and hair.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cosmetic or dermatological composition containing spider silk protein, for hair or skin care, in make-up or sunscreens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  528 GYGGLGSQG-AGRGGLGGQGAGAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGG 586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----AAAAAAGGAGQGGYGCLGSQCT--SGRGGLGGQGA-- 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----AAAAAAGGAGGGGYGGLGS 49
                                                                                                                                                              Polymer of an analogue of spider silk protein spidroine major 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 1036; DB 20; Length 606;
Pred. No. 5.6e-73;
4; Mismatches 23; Indels 102;
                                                                                                                                                                                Spider silk protein; spidroine major 1; cosmetic; make-up; dermatological compositions; hair care; skin care; sunscreen; hormone; moisturizer; skin disorder; skin disorder.
                                                                                                                                                                                                                                                                                                                                                                                                              Garson JC, Arraudeau JP,
                                                                                                                                                                                                                                                                      1..101
/note= "monomer unit"
                                                                                                                                                                                                                                                            Location/Qualifiers
                                                                                              AAY40100 standard; protein; 606 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 8; Fig 4B; 32pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  69.88;
                                                                                                                                                                                                                                                                                                                                               98FR-0001614.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64.98;
                                                                                                                                                                                                                                                                                                                                                                    98FR-0001614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 GSMASGRGGLGGQGAGAAA-
                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 239; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1999-510729/43.
                                                                                                                                                                                                                                                                                                                                                                                        (OREA ) L'OREAL SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            606 AA;
                      272 YGGLGSQG 279
                                                                                                                                                                                                                          Synthetic.
Nephila clavipes.
                                          587 YGGLGSQG 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               105 AA-----
                                                                                                                                                                                                                                                                                                     FR2774588-A1
                                                                                                                                                                                                                                                                                                                                               11-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                   11-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                             Philippe M,
                                                                                                                                        19-NOV-1999
                                                                                                                                                                                                                                                                                                                          13-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                  AAY40100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                             Key
Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                          RESULT 6
                                                                                     AAY40100
 ð
                      ð
                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged obligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (11) as adhesion surfaces for culturing cells; and (11) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity).

Also they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and purified. The modular construction of the invention facilitates and construction of the invention facilitates.

Characteristication or modulate solubility. This sequence represents the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New DNA encoding synthetic splder silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                   -----AAAAAAGGAGQGGYGGGGGT--SGRGGLGGQGA---GAAAAAAAAGGAGQG
GYGGLGSQGTSGRGGLGGQGAGAAAA---------AAAAAAGGAGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Spidroin; spider; silk; fibre; film; membrane; wound; filter; SO1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N. clavipes spidroin synthetic homologue SO1 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Guehrs K;
```

9

Page

```
Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                 221 -----
                                          proteins
                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAM50037;
                                                                                                                                                                                                                                                                                                                                                                                                                            259
                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAM50037
  쉽
                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        XX
                                               13;
                                                                                                                                                                          220
                                                                     LGGQGAGAAAA-------AAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 135
                                                                                                                           136 GQGAGAAAAAAAAA.----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA---- 182
                                                                                                                                                            311
                                                                                                                                                                                Gaps
                                                             ------AAAAAGGAGGGY 44
                                                                                       45 GGLGSQGTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGG 95
                                                                                                                                                     ---AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAA--
                                                                                                                                                                                                                                                                                                                                  Spidroin; spider; silk; fibre; film; membrane; wound; filter; SO1
synthetic N. clavipes spidroin-1 homologue SO1 described in the
                                 DB 23; Length 630;
                                               12; Indels 145;
                                                                                                                                                                                                                                                                                                                                                                                         "Synthetic spidroin homologue SO1"
                                                                                                                                                                                                                                                                                                                    N. clavipes spidroin synthetic homologue SO1 protein #2.
                                                                                                                                                                                                                              430 GGAGQGGYGGLGSQGAGRGGQCAGAAAAAGGAGQGGYGGLGSQG 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                         Pred. No. 1.3e-72
                                                                                                                                                                                                                                                                                                                                                                                                           673..676 "fretention signal"
                                                                                                                                                                                                                                                                                                                                                                          /label- LeB4_signal_peptide
                                 Score 1031.5;
                                               6; Mismatches
                                                                                                                                                                                                                                                                                                                                                              .ocation/Qualiflers
                                                                                                                                                                                                                                                                            AAM50047 standard; Protein; 676 AA
                                                            11 GSMASGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                      "c-myc-tag"
                                 69.5%;
59.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-MAR-2001; 2001DE-1013781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09-JUN-2000; 2000DE-1028212
24-OCT-2000; 2000DE-1053478
                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                         /note= "Sy
660..672
/note= "c
673..676
                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                  29..659
                                       Similarity
                    630 AA;
                                                                                                                                                                                                                              259 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                 DE10113781-A1
                                                                                                                                                                                                                                                                                                        18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                              13-DEC-2001
                                              242;
      invention
                                                                                                                                                                                                                                                                                                                                               Synthetic
                    Sequence
                                                                                                                                                                                                                                                                                          AAM50047;
                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                   Pept1de
                                                                                                                                                                                                                                                                                                                                                                                  Protein
                                                                                                                                                                                                                                                                                                                                                                                               Region
                                                                                                                                                                                                                                                                                                                                                                                                             Domain
                                              Matches
                                                                                                                  96
                                                                                                                               197
                                                                                                                                                                                                                                                                                                                                                              Key
                                                                                                                                                                                                                                                               RESULT
                                                                                      ö
                                                                                                    셤
88333
                                                                         셤
                                                                                                                  ð
                                                                                                                              g
                                                                                                                                             ò
                                                                                                                                                          엽
                                                                                                                                                                                    g
                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                合
                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                          ð
```

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and artificial organs; (ii) as adhesion surfaces for culturing cells; and (iii) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to also they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and composed of the modulate solubility. This sequences, e.g. to simplify purification or modulate solubility. This sequence represents a construct composed of the Leak signal peptide—encoding sequences; e.g. to simplify themologue sol, a c-Myc-tag and an endoplasmic reticulum (ER)-retention signal described in the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13;
                                                                                                                                    New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 136 GQGAGAAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA---- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             340 AGQGGLGGQGAGQGAQAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAAGGAG 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA-- 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 1031.5; DB 23; Length 676;
Pred. No. 1.4e-72;
6; Mismatches 12; Indels 145; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 107 GSQGAGRGGIGGQGAGAAAAAGGAGQGGYGGIGSQGAGRGGQGAGAAAAAAGGAGQGGY 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----AAAAAGGAGQGGYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     45 GCLGSQGTSGRGCLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------AAAAAAGGAGGGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      183 ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------AAAAAGGAGGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                458 GGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGGGYGGLGSQG 502
   Guehrs K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ź
Grosse F,
                                                                                                                                                                                                                                                                        Example 1; Fig 10A; 88pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAM50037 standard; Protein; 1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             69.5%;
59.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-SEP-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 GSMASGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96 LGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 59.8
Matches 242; Conservative
Conrad U,
                                                                WPI; 2002-123561/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         676 AA;
```

```
g
                                                                         ò
                                                                                                 Q
                        à
                                                                                                                                                                                                                                                                                                                                                                     spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at clear 1848 homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and (iii) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to close with excellent mechanical properties (strength and elasticity). Choso they retain water solubility after long-term boiling in aqueous collutions and since they are also soluble in organic solvents but contributed at high saft concentration, they are easily extracted and purified. The modular construction of the invention facilitates construction of the invention facilitates to incorporation or modulate solubility. This sequences, e.g. to simplify synthetic Northern and sequence represents the construction.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13;
                                                                                                                                                                                                                                                                                      encoding synthetic spider silk protein, useful e.g. for closing comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GQGAGAAAAAAAAAA-----AGGAGQGYGGYGGLGSQGTSGRGLGGQGAGAAAA---- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                704 GSQGAGRGGLGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGY 763
                                    Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOISOI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGLGSQGTSGRGGLGGQGAGAAAAAAAAAA------AGGAGQGGYGGLGSQGTSGRGG 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----AAAAAAGGAGGGY 44
                                                                                                                                                                                                                                                                                                                                                             This invention describes a novel DNA sequence, encoding a synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -AAAAAAGGAGGGGGGGGGGGGGGGGGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 1031.5; DB 23; Length 1255; Pred. No. 2.3e-72; 6; Mismatches 12; Indels 145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N. clavipes spidroin synthetic homologue SO1SO1 protein.
                                                                                                                                                                                                 (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                       Conrad U, Grosse F, Guehrs K;
                                                                                                                                                                                                                                                                                                                                    Claim 22; Page 29-34; 88pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         69.58;
59.88;
                                                                                                                                    21-MAR-2001; 2001DE-1013781.
                                                                                                                                                           2000DE-1028212.
2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GSMASGRGGLGGQGAGAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 59.8 Matches 242; Conservative
                                                                                                                                                                                                                                               WPI; 2002-123561/17.
N-PSDB; ABL61038.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1255 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                    DE10113781-A1
                                                                                                                                                           09-JUN-2000;
24-OCT-2000;
                                                                                                                                                                                                                       Scheller J,
                                                                                                            13-DEC-2001
                                                            Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          nvention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                              proteins
                                                                                                                                                                                                                                                                                     New DNA (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           183
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
```

```
spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidonin protein. The synthetic protein has at least 84% homology with the Nephila claylpes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (1i) as adhesion surfaces for culturing cells; and (1ii) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity).

Also they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high saft concentration, they are easily extracted and purified. The modular construction of the invention facilitates incorporation of additional peptide-encoding sequences, e.g. to simplify purification or modulate solubility. This sequence represents the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOISOLSOL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA encoding synthetic spider silk protein, useful e.g. for closing ids, comprises modules that encode repeating units of spidroin
                                                                                               Gapa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        encoding a synthetic
AGQGGLGGQGAGAGAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAAAGGAG
                                                                  --------AAAAAGAGQGGYGGLGSQGTSGRGLGGQGAAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 1031.5; DB 23; Length 1880;
Pred. No. 3.1e-72;
6; Mismatches 12; Indels 145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N. clavipes spidroin synthetic homologue SO1SO1SO1 protein.
                                                                                                                                                                                                                     1055 GGAGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQG 1099
                                                                                                                                                                                     -----AAAAAAGGAGGGKGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This invention describes a novel DNA sequence,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Guehrs K;
                                                                                                                                                                                                                                                                                                                                                                                                 AAM50039 standard; Protein; 1880 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 22; Page 38-45; 88pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Grosse F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     69.5%;
59.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21-MAR-2001; 2001DE-1013781.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000DE-1028212.
2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2002-123561/17.
N-PSDB; ABL61040.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 242; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DE10113781-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09-JUN-2000;
24-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-DEC-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAM50039;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                     259
                                                            221
                                                                                                                                                                                                                                                                                                                                       RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New 1
                                                                                                                                                                                                                                                                                                                                                                       AAM50039
```

œ

```
A process has been developed for the production of a DNA fragment encoding silk protein. The process involves: (a) selecting target DNA, from a silk producing spider, that contains many repetitive and non-repetitive regions; (b) selecting a single-stranded DNA primer of at least 10 nucleotides with a sequence that is complementary to a region of the target; (c) repetitively combining the primer with melted target
                                                                                                                                      252
                                                                                                                                                            136 GQGAGAAAAAAAAA.-----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA---- 182
                                                                                                                                                                                                                                 138
                                                                   --- AAAAAAGGAGGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                      95
                                                                                                                                197 LGGQGAGAAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRG
                                                                                                                                                                               - AAAAAAGGAGOGGY
                GGLGSQGTSGRGGLGGQGAGAAAAAAAA.-----AGGAGQGGYGGLGSQGTSGRGG
                                                                                                                                                                                                                                                                         ---AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA--
                                                                                                                                                                                                                  ------AAAAAAGAGQGGYGGLGSQCTSGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              High strength film; fibre; woven article; parachutes; sails; absorber; body armour; heavy metal; biological weapon; chemical; flavour; fragrance; Nephila clavipes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New opt. multimerised DNA sequences encoding spider silk pro contg. both repetitive and non-repetitive sequences, useful making high strength films, fibres, woven articles etc.
                                                                                                                                                                                                                                                                                                                          259 -------AAAAAAGGAGGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                           430 GGAGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQG 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nephila clavipes spider silk protein.
                                                                                                                                                                                                                                                                                                                                                                                                                      AAW27178 standard; Protein; 646 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Example 2; Fig 1; 57pp; English,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . 96WO-US13767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95US-0517694.
11 GSMASGRGGLGGQGAGAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Elion GR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1997-179272/16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BASE/) BASEL R M. (ELIO/) ELION G R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAT85356.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO9708315-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22-AUG-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-DEC-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          06-MAR-1997
                                                                                                                                                                                                                                                                      221 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Basel RM,
                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW27178;
                                                     45
                                                                                                        96
                                                                                                                                                                                                                  183
                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                           AAW2717
                                                                                                                                                                                    g
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                            셤
                                                                                                        ò
                                                                                                                                q
                                                                                                                                                            ò
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                          ò
                                                     ð
```

```
13;
DNA, incubating the mixture with nucleotides and a DNA polymerase with profireading activity to produce a DNA fragment which is complementary to the target and is at least 2 kb long. The present sequence encodes the spider silk protein from Nephila clavipes. The DNA fragment can be used to make fibres, films, woven articles, e.g. for use in parachutes, sails, body armour, and absorbers (e.g. of heavy metals, biological weapons, DNA, chemicals, flavours and fragrances). The high molecular weight (90-250 kD) of spider silk proteins can be produced on a commercial scale (at over 2 g/l cell mass). It has better tensile strength and elasticity than silkworm silk. Inclusion of both repetitive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      110 AAAGGAGQGGYGGLGSQGTSGRGGGGGGAGAAAAAAAG------GAGQGGYG 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            402 GYGGLGGQGAQQGGYGGLGSQG-SGRGGLGGQGAGAAAAAAGGAGGGLGGQGAGAA 460
                                                                                                                                                                                                                                                                                                     110 OGAGRGGOGAGAAAAAAGGAGOGGYGGOGAGOGGYGGLGSOGAGRGGLGGOGAGAAAAA 169
                                                                                                                                                                                                                                                                                                                                                                                                          343 LGNQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGGGGAGAAAAAGGAGQG 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                 11 GSMASGRGGLGGQGAGAAAA-------AAAAAAGGAGQGGYGGLGS 49
                                                                                                                                                                                                                                                                                                                                                                                                                            170 AGGAGQGGLGGQGAGGAGAAAAAAGGAGQGGYGGLGSQG----AAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------SGRGCLGGQGAGAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 160 GLGSQGTSGRGGLGGQGAGAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----AGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAA------A
                                                                                                                                                                                                                                                      Indels 180;
                                                                                                                                                            and non-repetitive regions ensures isolation of stable clones
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Spider silk protein; spidroine major 1; cosmetic; make-up; dermatological compositions; hair care; skin care; sunscreen; hormone; moisturizer; skin disorder;
                                                                                                                                                                                                                         Length 646;
                                                                                                                                                                                                                    Score 1031; DB 18;
Pred. No. 1.4e-72;
6; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Spider silk protein spidroine major 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY40097 standard; protein; 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       260 AAAAAGGAGQGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         461 AAAAAGGAGQGGYGGLGSQG 480
                                                                                                                                                                                                                      69.4%;
55.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98FR-0001614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                      Best Local Similarity 55.5
Matches 244; Conservative
                                                                                                                                                                                         646 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nephila clavipes
                                                                                                                                                                                                                                                                                                                                             50 QGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FR2774588-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY40097;
                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                          69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY40097
    8888888888888888
                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
```

σ

```
Nephila clavipes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
AAW53346
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                  13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Orb-weaving spider; spidroin 1; silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation.
                                                                                      Cosmetic or dermatological composition containing spider silk protein, for hair or skin care, in make-up or sunscreens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGYGGLGSQGAGRGGQGAGAAAAAAGGAGGRGYGCLGNQG-AGRGGLGGQGAGAAAAAAA 443
                                                                                                                                                                                                                                                                                                                                                                                                        153 GLGSQG-AGRGGLGGQGAGAAAAAGGAGQGGYGGLGGQGAQQGGYGGLGSQG-AGRGGL 210
                                                                                                                                                                                                                                                                                                                                                                                                                              GGQGAGAAAAAA-------AAAAGGAGGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                          384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93 GSQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGNQGAGRGGQGAAAAAAGGAGQGGYG 152
                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                   96
                                                                                                                                                                                                                                                                                                                                        -----AAAAAAGGAGGGYG 45
                                                                                                                                           The present sequence represents the natural spider silk protein spidroine major 1. The protein improves the moisturizing/softening action of the compositions. The protein, and its fragments are used in cosmetic or dermatological compositions. These compositions have use as hair or skin care products; and make-up or sunscreens. As the protein is a good, persistent film-formers on the skin of low surface density, it can be used for delivery of active agents that are generally difficult to administer, e.g. vitamins, hormones, moisturizers or agents for treating disorders of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 936 QGGLGGQGAQQGAGAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAVAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                   GLGSQGTSGRGGLGGQGAGAAAAAAAAAA...-.AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         136 GQGAGAAAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA----
                                                                                                                                                                                                                                                                                                                  Indels 142;
                                                                                                                                                                                                                                                                                           DB 20; Length 651;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------AAAAAGGAGQGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                         68.4%; Score 1015; DB 20;
59.2%; Pred. No. 2.5e-71;
iive 8; Mismatches 14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Spider natural silk protein Spidroin 1.
                                             Arraudeau JP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAU11781 standard; protein; 651 AA
                                                                                                                                                                                                                                                                                                                                      11 GSMASGRGGLGGQGAGAAA-----
                                                                                                                      Claim 3; Fig 1; 32pp; French
98FR-0001614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                  Conservative
                                            Garson JC,
                                                                WPI; 1999-510729/43.
                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 238; Conserv
                      OREA ) L'OREAL SA
                                                                                                                                                                                                                                                                    651 AA;
                                                                                                                                                                                                                                               skin and hair.
11-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26-MAR-2002
                                           Philippe M,
                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAU11781;
                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    183
                                                                                                                                                                                                                                                                                                                                                                                  46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            260
                                                                                                                                                                                                                                                                                                                                                                                                                             97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAU11781
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a protein state of the expression of an SLP transgene and T is a proteins in green plants. The silks and silk-like proteins any broteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, flexible tie downs for certain electrical components, or as a blomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is orb-weaving spider spidrolni, a natural silk the method of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The invention relates to a method of producing {	t silk-like} proteins (SLP)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           136 GQGAGAAAAAAAA-----AGGAGQGGYGGLGSQCTSGRGGLGGQGAGAAA---- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93 GSQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGNQGAGRGGQGAAAAAAGGAGQGGYG 152
                                                                                                                                                                                                                                                                                                                                                                                                         Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein expression cassette to a green plant -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----- AAAAAAGGAGGGYG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGQGAGAAAAAA---------AAAAGGAGQGGYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               46 GLGSQGTSGRGGLGGQGAGAAAAAAAAA...--AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  267 GEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            183 ------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 651;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --------AAAAAGGAGQGCYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           444 GGAGQGGYGGLGNQGAGRGGQGAAAAAGGAGQGGYGGLGSQG 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 1015; DB 23
Pred. No. 2.5e-71;
8; Mismatches 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11 GSMASGRGGLGGQGAGAAA-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Page 69-71; 93pp; English.
                                                                                                                                                                                                                                    E I.
                                                                                                                                                                                                                                    ပ္ပ
                                                                                                                                                                                                                                        ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    68.4%;
59.2%;
                                                                                                                                                                          25-MAY-2000; 2000US-206968P
                                                                                                                    24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                    (DUPO ) DU PONT DE NEMOURS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 59.2
Matches 238; Conservative
                                                                                                                                                                                                                                                                                                                                                     WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             651 AA;
WO200190389-A2.
                                                          29-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                           Yang JG;
```

19

Dec

Thu

```
RV;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  68.4%;
59.2%;
                                                                                                                                                                                                                                                           91US-0684819,
90US-0511792,
                                                                                                                                                                                                                                        94US-0317844
                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11 GSMASGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                          Lewis
                                                                                                                                                                                                                                                                                                                                                                 the production of spider
                                                                                                                                                                                                                                                                                                                                                        vector and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGGGAGAAAAAA - - - -
                                                                                                                                                                                                                                                                                       (UYWY-) UNIV WYOMING
                                                                                                                                                                                                                                                                                                                          WPI; 2000-061225/05.
N-PSDB; AAZ38195.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                         Xu M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               718 AA;
                                                                                                                                                                                 Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                        Isolated DNA,
                                                                                                                                                                                                                                        04-0CT-1994;
                                                                                                                                                                                                                                                          15-APR-1991;
20-APR-1990;
                                                                                                                          08-MAR-2000
                                                                                                                                            N. clavipes
                                                                                                                                                               Spider silk
                                                                                                                                                                                                   JS5989894-A
                                                                                                                                                                                                                     23-NOV-1999
                                                                                                                                                                                                                                                                                                         Hinman MB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 238;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
                                                                                                       AAY59070;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                     260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     136
                                                                   RESULT 15
                                                                            AAY 59070
                                                                                                         g
                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13;
                                                                                                                                                                                                                                                                                                                                The present sequence represents a spider silk protein from the present invention. Spider silk proteins, and peptide fragments within the proteins, can be produced and purified independently and can then be mixed and made into fibres that have higher tensile strengths and elasticity than naturally occurring fibres. The fibres can be used in mixed composites. The invention allows the two naturally occurring Nephila clavipes silk proteins to be produced independently so that they can later be combined to form silk fibres of high tensile strength
                                                                        silk protein; tandem repeat; fibre; dragline; elasticity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA-- 259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                136 GQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA--- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----AAAAAGGAGGGYG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLGSQGTSGRGGLGGQGAGAAAAAAAAA.-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGQGAGAAAAAA------------AAAAGGAGQGGYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  326 QGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAVAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----AAAAAAGGAGGYGGLGSQGTSGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 1015; DB 19; Length 718;
Pred. No. 2.7e-71;
3; Mismatches 14; Indels 142
                                                                                                                                                                                                                                                                                            Recombinant spider silk proteins - useful for making fibres
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GSMASGRGGLGGQGAGAAA-------
                                                       Nephila clavipes spider silk protein
                                                                                                                                                                                                                                                                                                              Claim 1; Column 29-34; 68pp; English
Ş
AAW53346 standard; Protein; 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8;
                                                                                                                                                                                        90US-0511792.
94US-0317844.
95US-0425069.
                                                                                                                                                                                                                                                Xu M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                68.48;
                                                                                                                                                            95us-0425069
                                                                                                                                                                              91US-0684819
                                                                                                                                                                                                                                                                                                                                                                                                                                                        59.28;
                                                                         clavipes;
strength;
                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                               Hinman MB, Lewis RV,
                                                                                                                                                                                                                            (UYWY-) UNIV WYOMING
                                                                                                                                                                                                                                                               WPI; 1998-270437/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                              718 AA;
                                                                         Nephila c
                                                                                                   Nephila clavipes
                                                                                                                                                                                                                                                                           N-PSDB; AAV23249
                                                                                                                                                                                                                                                                                                                                                                                                            elasticity.
                                                                                                                                                            19-APR-1995;
                                                                                                                                                                                        20-APR-1990;
04-OCT-1994;
                                    06-JUL-1998
                                                                                                                                                                              15-APR-1991;
                                                                                                                                                                                                          19-APR-1995;
                                                                                                                       US5728810-A.
                                                                                                                                          17-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  238;
                  AAW53346;
                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                         Spider;
                                                                                   coccon;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             221
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
protein; dragline silk protein; major ampullate gland; fiber.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The invention provides isolated CDNA molecules coding for spider silk proteins. The spider silk proteins are characterized by repeating alpha and beta regions and optional variable regions. The DNA sequences are useful in the production of spider silk protein by recombinant DNA techniques. The recombinant spider silk proteins may be used for the production of fibers. The present sequence represents the spider silk protein 1, derived from the major ampullate gland of Nephila clavipes.
385 GGYGGLGSQGAGRGGQGAGAAAAAGGAGQRGYGGLGNQG-AGRGGLGGQGAGAAAAAA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cell encoding for and useful in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----AAAAGGAGQGGYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21,
                                                                      Score 1015; DB 21
Pred. No. 2.7e-71;
3; Mismatches 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          silk protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  transformed
                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                           spider silk protein 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 1; Fig 6A-D; 65pp; English.
                                                                                                                                                                                                        AAY59070 standard; Protein; 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

This Page Blank (uspto)

```
19, Appl
1, Appl
1, Appl
2, Appl
3, Appl
8, Appl
6, Appl
1, Appl
1, Appl
1, Appl
1, Appl
4, Appl
4, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appl
Appl
Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 21, Appli
Sequence 4, Appli
                                                                                                            December 18, 2002, 16:14:14; Search time 6.69992 Seconds (without alignments) 1277.936 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1 Sequence 2 Sequence 3 Sequence 3 Sequence 4 Sequence 4 Sequence 5 Sequence 5 Sequence 5 Sequence 6 Sequence 6 Sequence 6 Sequence 6 Sequence 7 Sequence 6 Sequence 6 Sequence 7 Sequence 6 Sequence 7 
                                                                                                                                                                                         US-09-490-291-2
1485
1 MRGSHHHHHGSMASGRGGL.....YGGLGSQGTSGIRRPAAKLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Issued_patents_AA:*

1: /cgn2_6/ptodatca1/iaa/5A_COMB.pep:*

2: /cgn2_6/ptodatca1/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodatca1/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodatca1/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodatca1/iaa/pcruS_COMB.pep:*

6: /cgn2_6/ptodatca1/iaa/pcruS_COMB.pep:*
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-556-978B-21
US-09-247-806-4
US-09-247-806-1
US-08-425-069-2
US-08-317-844B-2
US-08-317-844B-2
US-08-356-978B-63
US-08-556-978B-63
US-08-556-978B-63
US-09-247-806-8
US-09-247-806-8
US-09-247-806-8
US-09-247-806-2
US-08-425-190B-1
US-08-425-069-4
US-08-425-069-4
US-08-556-978B-61
                                                                                                                                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-477-509B-64
US-08-707-237A-35
                                                                                                                                                                                                                                                                                                                                  262574 seqs, 29422922 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                           - protein search, using sw model
                                                                                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1036
1036
1015
1015
1015
1015
1015
984
976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             674
610.5
610.5
593.5
553.5
554
554
520
520
520
520
520
520
                                                                                                                                                                                                                                                                         Scoring table:
                                                                         OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Database :
                                                                                                                                                                                                                                                                                                                                    Searched:
                                                                                                                  Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š.
```

Appl Appl Appl Appl Appl Appl Appl Appl		13;	
964, 931, 931, 94, 94, 93, 93, 936,		Gaps	GS 49
Sednence Sed		Length 606; Indels 102;	-AAAAAAGGAGGGYGGLGS
	Q AN	Length Indels	VAAAAGG
S-08-482-085B-64 S-08-475-411A-31 S-08-478-029A-31 S-08-175-155-48 S-08-707-237A-54 S-08-707-237A-54 S-08-86-029-10 S-08-477-509B-83 S-08-447-791A-83 S-09-444-791A-83 S-09-444-791A-83 S-09-447-411A-41 S-08-475-411A-41 S-08-475-411A-36 S-08-475-411A-36 S-08-478-029A-41 S-08-478-029A-41 S-08-478-029A-41 S-08-478-029A-31 S-08-478-029A-36 S-08-478-11A-36	ALIGNMENTS I US/08556978B IX, STEPHEN F. NOVEL RECOMBINANTLY PRODUCED SPIDER SILK ANALOGS 107 SS: 107 SS: 108 NUT DE NEMOURS AND COMPANY STREET OF AMERICA MICROSOFT WINDOWS 95 TH WORD FOR WINDOWS 95 TH WO	Score 1036; DB 4; Pred. No. 5.7e-77; ; Mismatches 23;	W
0.5 - 0.5 -	ALIG SY08556978B SY08556978B STEPHEN F. VEL RECOMBIN UDER SILK AN USTREET CROSOFT WIND WORD FOR WIND WORD FOR WIND USY08/556,97 US/08/556,97 US/08/556,97 US/08/556,97 US/08/556,97 US/08/556,97 US/08/556,97 US/08/556,97	8; Score 8; Pred 4; M	
11177 11177 11177 11059 11059 11001 11101 11332 1332 1332 1332 1338 1038 1038	ALIGNMEN 1.1.cation US/08556978B 6.9 TION: TION: TATION: TATION TATION: TATION TATION: TATION TATION: TATION TATION: TATION	69.8 64.9 vative	GAGAAAA
0.000 444444444444444444444444444444444	dugeningo ozioania a Hzarezzazzote weroca a	Similarity 9; Conser	GSMASGRGGLGGQGAGAAAA :
520 520 520 520 507 507 507 507 488 488 480.5 476.5 476.5		atch cal 23	11 GSMAS
00000000000000000000000000000000000000	RESULT 1 US-08-556-6 Sequence Patent NG GENERAL TITLE TITLE TITLE TITLE TITLE CORREG CORREG COMPUT C	Query M Best Lo Matches	ò

271

```
GQGAGAAAAAAAAA -----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA---- 182
   468 GQGAGAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQG 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---AAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 1015; DB 4; Length 651;
Pred. No. 3e-75;
8; Mismatches 14; Indels 142;
                                                                                                                                                                                                                                                                                                                                                             E: E. I. DU PONT DE NEMOURS AND COMPANY 1007 MARKET STREET
                                                                                                                                                                                                                                                                          APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
TITLE OF INVENTION: SPIDER SILK ANALOGS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
OFFRATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,600
FILING DATE: JUNE 15,1933
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: GR-9389-A
TELECHONE: 302-892-8112
TELEPHONE: 302-773-0164
                                                                                                                                                                                      RESULT 3
US-08-556-978B-19
US-08-556-978B-19
Sequence 19, Application US/08556978B
Patent No. 6268169
GENERAL INFORMATION
                                     233 GYGGLGSQGTSGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : 19:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   68.4%;
59.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 651 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11 GSMASGRGGLGGQGAGAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 68.4 Best Local Similarity 59.2 Matches 238; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: unknown MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                 STREET: 1007 MAR
CITY: WILMINTON
STATE: DELAWARE
                                                                                                     272 YGGLGSQG 279
                                                                                                                                      587 YGGLGSQG 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-556-978B-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          136
                                     ò
                                                                   g
                                                                                                      ò
                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: ARRAUDEAU, Jean-Plerre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER APPLICATION NUMBER: FR 98/01614
EARLIER APPLICATION NUMBER: FR 98/01614
EARLIER PILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOSTWARE: PATENTIN Ver. 2.1
230 GSQGAGRGGLGGQGAGAAAAAAAGGGGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGS 289
                                                   290 QGAGQGGYGGLGSQGAGRGGGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGGGGAGAA 348
                                                                                                     ------AAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGA-- 139
                                                                                                                      183 ----AAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGA---GAAAAAAAAAAGGAGGG 232
                                                                                                                                                                                                                                                                                                                           GYGGLGSQGTSGRGGLGGQGAGAAA--------AAAAAGGAGQGG 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----AAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGA---GAAAAAAAAAAAGAGG 232
                               50 QGT--SGRGGLGGQGA---GAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              50 QGT--SGRGGLGGQGA---GAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 105 AA----SGRGGLGGQGA-- 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----AAAAAGGAGGGGGGGGG 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23; Indels 102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 1036; DB 4; Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 5.7e-77
4; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 4, Application US/09247806
Patent No. 6280747
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          69.8%;
64.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11 GSMASGRGGLGGQGAGAAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
ORGANISM: Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 64.9
Matches 239; Conservative
                                                                                                                                                                                                                                                                                                                                                                                               1111111
587 YGGLGSOG 594
                                                                                                   105 AA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q ID NO 4
LENGTH: 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-247-806-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-247-806-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                          233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      183
                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                      셤
                                                                 셤
                                                                                                 ö
                                                                                                                                 a
                                                                                                                                                                     ð
                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
```

13;

Gaps

```
Sequence 2, Application US/08425069
Patent No. 5728810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 718 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 238; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; MOLECULE TYPE: protein US-08-425-069-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                        U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION:
                                                                                                                                                                                                                                                   22046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                   COUNTRY:
   US-08-425-069-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13;
                                                                                                                                                                                                                                                                                         APPLICANT: PHILLIPPE, Michel
APPLICANT: PRILLIPPE, Michel
APPLICANT: GRESON, Jean-Claude
APPLICANT: ARRAUDEAU, Jean-Pletre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
SARLIER FILING DATE: 1999-02-11
SOFTWARE: PATCHIN VOIC: 2.1
SOFTWARE: PATCHIN VOIC: 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              885 GGYGGLGSQGAGRGGQGAGAAAAAGGAGQRGYGGLGNQG-AGRGGLGGQGAGAAAAAA 443
267 GEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAGGAG 325
                                                                                                                 385 GGYGGLGSQGAGRGGQGAGAAAAAGGAGQRGYGGLGNQG-AGRGGLGGQGAGAAAAAA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GOGAGAAAAAAAAAA.-----AGGAGOGGYGGLGSQGTSGRGGLGGQGAGAAA---- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    326 QGGLGGQGAGQGAGAGAAAAAAGGAGQGGYGGLGSQC-AGRGGLGGQGAGGAAAAAAAAAGAGQ 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA-- 259
                                                           384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   46 GLGSQGTSGRGGLGGQGAGAAAAAAAAA------AGGAGQGGYGGLGSQGTSGRGGL 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11 GSMASGRGGLGGQGAGAAAA--------AAAAAAGGAGGGYG 45
                                            326 QGGLGGQGAGQGAGAAAAAAGGAQGGYGGGGGGAGAAAAAAAGGAGQ
                                                                                      -----AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14; Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 4; Length 651;
                                                                                                                                                               -----AAAAAGGAGQGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGAGOGGYGGLGNQGAGRGGOGAAAAAGGAGQGGYGGLGSQG 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. 3e-75;
8; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 1015;
                                                                                                                                                                                                                                                Sequence 1, Application US/09247806
Patent No. 6280747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          68.4%;
59.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
ORGANISM: Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 59.2 Matches 238; Conservative
                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
                                                                                                                                                                                                                                   US-09-247-806-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-247-806-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
 a
                              õ
                                                      q
                                                                                     ò
                                                                                                                                              ð
                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
```

```
13;
                                                                                              TRANSFORMED CELL
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Xu, Ming
APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CEI
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----AAAAGGAGQGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           136 GQGAGAAAAAAAAA.....AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA.... 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11 GSMASGRGGLGGQGAGAAAA------AAAAAAAGGAGQGGYG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLGSQGTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             326 QGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAVAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14; Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------AAAAAGGGGGGGGGGGGGGGGGGGGGGGGGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 718;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              - AAAAAGGAGQGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
ELING DATE: US/08/425,069
FILING DATE: 19-APR-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 1015; DB 1;
Pred. No. 3.3e-75;
8; Mismatches 14;
                                                                                                                                           AUDRESSEE: Birch, Stewart, Kolasch & Birch STREET: 301 No. 5728810th Washington Street CITY: Falls Church STATE: Virgin's
                                                                                                                                                                                                                                                                                            ZIP: 22046
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 14
TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           68.48;
59.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEPHONE: (703) 205-8000 TELEFAX: (703) 205-8050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGGGAGAAAAA
```

```
385 GGYGGLGSQGAGRGGQGAGAAAAAGGAGQRGYGGLGNQG-AGRGGLGGQGAGAAAAAA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----AGGAGOGGYGGLGSQCTSGRGGLGGQGAGAA---- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 142; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   - AAAAAAGGAGQGGYG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        46 GLGSQGTSGRGGLGGQGAGAAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 747;
                                                                           444 GGAGQGGYGGLGNQGAGRGGQGAAAAAGGAGQGGYGGLGSQG 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 1015; DB 3;
Pred. No. 3.4e-75;
8; Mismatches 14;
                                                                                                                                                                                                         APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HOMAN FIBROUS PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: FastSEQ for Windows Version 2.0 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/034,177
FILING DATE: HERRWITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF-0486 US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US/09/034,177
HEREWITH
                                                                                                                                                                  Sequence 3, Application US/09034177 Patent No. 6127146 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: B1111049, Lucy J
RECISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0.
TELEPHONE: 650-855-0555
TELEPHONE: 650-845-4166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           68.4%;
59.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS: LENGTH: 747 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11 GSMASGRGGLGGQGAGAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best_Local Similarity 59.2
Matches 238; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                       ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MMEDIA...
LIBRARY: Genda...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    136 GOGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IMMEDIATE SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; CLONE: C
US-09-034-177-3
                                                                                                                                                 US-09-034-177-3
                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                       260
                                                                                                                                RESULT 7
                                                                       셤
                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13;
                                                                                                                                APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Xu, Ming
APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 NO. 5989894th Washington Street
CITY: Falls Church
STATE: Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97 GGQGAGAAAAAA-------AAAAGGAGQGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                136 GQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA--- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -AAAAAAGGGGGGYGGLGSQCTSGRGGLGCQGAGAAAA-- 259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           326 QGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAVAAAAAGGAGQ 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11 GSMASGRGGLGGQGAGAAAA------AAAAAAGGAGQGGYG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     46 GLGSQGTSGRGGLGGQGAGAAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          183 ------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14; Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 1015; DB 2; Length 718;
Pred. No. 3.3e-75;
444 GGAGQGGYGGLGNQGAGRGGQGAAAAAGGAGQGGYGGLGSQG 485
                                                                                                                                                                                                                                                                                                                                                                             ZIP: 22046

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
US/08/317,844B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US/08/317,844B
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
                                                                                  Sequence 2, Application US/08317844B Patent No. 5989894 GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        68.4%;
59.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                718 amino acids amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEX: 246345
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 238; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                              COUNTRY: U.S.A. ZIP: 22046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                       US-08-317-844B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-317-844B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                221
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
```

```
183 ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGG 228
                                                                                                                                                                                                 229 AGGGYGGLGSQGT------SGRGGLGGQGAGAAAA------- 258
                                                                                                                                                                                                                                      494 AGGGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGGGLGSQGAGQ 553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       213 GSQGAGRGGLGGQGAGAAAAAAGGAGQGGLGSQGAGQGAGAAAAAGGAGQGGYGGLGS 272
  330 AAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG- 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.1 GSMASGRGGLGGQGAGAAAA-------AAAAAAGGAGGGYGGLGS 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---SGRGGLG 97
                                                               442 QGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AAAAAAAGG
                                          142 AAAAAAAAGGGGGGGGGGGGGGT-----SGRGGLGGGGGAAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels 144;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 4; Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             50 QGTSGRGGLGGQGAGAAAAAAAAAGGAGGGYGGLGSQGT------
                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 233, Application US/08556978B
Patent No. 6268169
GENERAL INFORMATION:
APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
TITLE OF INVENTION: SPIDER SILK ANALOGS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 976; DB 4;
Pred. No. 4e-72;
4; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CR-9389-A
                                                                                                                                                                                                                                                                             08/077,600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,60
FILING DATE: JUNE 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: FLOYE, LINDA AXAMETHY
RECISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: GR-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: 302-773-0164 INFORMATION FOR SEQ ID NO: 23:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    65.7%;
59.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                302-892-8112
302-773-0164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               606 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET: 1007 MAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STATE: DELAWARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                             US-08-556-978B-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-556-978B-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                        ò
                                                                           qq
                                                                                                                  ò
                                                                                                                                                         g
                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGA 141
267 GEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAGGAG 325
                                      183 ------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA----- 220
                                                           326 QGGLGGQGACQGAGAAAAAGGACQGGYGGLGSQG-AGRGGLGGQGAGQAAAAAAAGGAGQ 384
                                                                                                                    -----AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAA-- 259
                                                                                                                                                       385 GGYGGLGSQGAGRGGQGAGAAAAAGGAGQRGYGGLGNQG-AGRGGLGGQGAGAAAAAA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              50 QGTSGRGGLGGQGAGAAAAAAAAAAAG-----GAGQGGYGGLGSQGTSGRGGLGGQGAG 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         215 GSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGS 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13; Indels 136; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----AAAAAAGGAGQGGYGGLGS 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 604;
                                                                                                                                                                                                                    444 GGAGQGGYGGLGNQGAGRGGQGAAAAAGGAGQGGYGGLGSQG 485
                                                                                                                                                                                               -----AAAAAGGAGGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5: E. I. DU PONT DE NEMOURS AND COMPANY
1007 MARKET STREET
                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
TITLE OF INVENTION: SPIDER SILK ANALOGS
NUMBER OF SEQUENCES: 107'
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 984; DB 4;
Pred. No. 9.1e-73;
4; Mismatches 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REGISTRATION NUMBER: 33,692
REPERENCE/DOCKET NUMBER: CR-9389-A
TELECOMMUNICATION:
TELEPHONE: 302-892-8112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CAMPUTER READSA

MEDIUM TYPE: DISKETTE, 3.50 INCH

COMPUTER: IBM PC COMPATIBLE

COMPUTER: IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                Sequence 63, Application US/08556978B Patent No. 6268169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08/011,600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,60
FILING DATE: JUNE 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA MARETHY
REGISTRATION NUMBER: 33,692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO: 63: SEQUENCE CHARACTERISTICS: LENGTH: 604 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66.3%;
60.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11 GSMASGRGGLGGQGAGAAA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 66.3°
Best Local Similarity 60.2
Matches 231; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   302-773-0164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-556-978B-63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: WILMINTON STATE: DELAWARE
                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
                                                                                                                                                                                             260 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19898
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE:
TELEFAX: 3
                                                                                                                                                                                                                                                                                                          US-08-556-978B-63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             103 AAAA-
a
                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                      ð
                                                                       g
                                                                                                              ŏ
                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

ø

```
489 AAAGGACQGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGS 548
                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 970; DB 4;
Pred. No. 1.2e-71;
4; Mismatches 12;
                                                         577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        259 -----AAAAAGGAGQGGYGGLGSQG 279
                                             549 QGAGQGAGAAAAAGGAGQGGYGGLGSQG
                                  -----AAAAAGGAGQGGYGGLGSQG
                                                                                                                            Sequence 6, Application US/09247806
Patent No. 6280747
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/08209747; Patent No. 573771 GENERAL INFORMATION: APPLICANT: Lewis, Randolph V. APPLICANT: CO19in, Mark TITLE OF INVENTION: CDNAS Encodi TITLE OF INVENTION: S11k Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                    65.3%;
58.9%;
                                                                                                                                                                                                                                                                                                                                                                               ; TYPE: PRT
; ORGANISM: Nephila clavipes
US-09-247-806-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98 GOGAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best_Local Similarity
Matches 229; Conserv
                                                                                                                                                                                                                                                                                                                                                                 909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 12
US-08-209-747-2
                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                               259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             183
    8
                                                         셤
                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13;
                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Pierre
THYLE OF INVENTION: CCSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6380-0355-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTIN VET. 2.1
                                                                                                                                                                                                                          548
                               136
                                             137 QGAGAAAAAAAAAGGGGGGGGGGGGGGT-----SGRGGLGGGGGGAGAAA-- 182
                                                                                                    224 AAAGGAGGGGYGGLGSQGT----- SGRGGLGGQGAGAAA------ 258
 273 OG----AGRGGOGAG----AAAAAAGGAGGGYGGLGSQGAGQGGYGGLGSQGAGRGGLG 324
                                                                                                                                       137 QGAGAAAAAAAAGAGGGGGGGGGGGGGT-----SGRGGLGGQGAGAAA-- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           437 AGGAGGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AAA 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98 GQGAGAAAA--------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               50 QGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLG 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11 GSMASGRGCLGGQGAGAAA-------AGAAAAAGGAGQGGYGGLGS 49
                                                                                                                                                                                                            ------AAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---SGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4e-72;
whes 11; Indels 144;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 4; Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 976;
                                                                                                                                                                                                                                                259 -----AAAAAGGAGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                               549 QGAGQGAGAAAAAGGAGQGGYGGLGSQG 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.
                                                                                                                                                                                                                                                                                                                                              Sequence 8, Application US/09247806
Patent No. 6280747
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     65.7%;
59.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAGGAGQGGYGGLGSQGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GQGAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
les 230; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 8
LENGTH: 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-247-806-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                           98
                                                                                                                                      183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    224
요
                                                g
                                                                                å
                                                                                                         셤
                                                                                                                                   á
                                                                                                                                                           g
                                                                                                                                                                                         õ
                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                            å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
13;
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: ARRAUDEAU, Jean-Pletre
TITLE OF INVENTION: COSMETIC OR DERWATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT APPLICATION NUMBER: US/09/241
EARLIER FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATCHIN VET. 2.1
SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 137 QGAGAAAAAAAAAGGAGGGYGGLGSQGT-----SGRGGLGGQGAGAAA-- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAA 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  439 AGGACQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AAA 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    215 GSQGAGRGGLGGQGAGAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGS 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 144; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11 GSMASGRGGLGGQGAGAAAA------AAAAAAGGAGQGGYGGLGS 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                50 QGTSGRGGLGGQGAGAAAAAAAAAAGAGGGGYGGLGSQGT-----SGRGGLG 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  224 AAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDNAs Encoding Minor Ampullate Spider
Silk Proteins
```

```
60 GQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAA 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAAGGAGGGYGGLGSQG-----TSGRGGLGGQGAGAAAAAAAAAAGGAGQG- 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -GYGGLGSQGT-----SGRGGLG-GQGAGAAAAA-----AAAAGGAGQGGYG 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15 SGRGGLGGQGAGAAAAAAAAAGGAGQGGYGG------LGSQGTSGRGGLG--- 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45.4%; Score 674; DB 1; Length 832; 46.4%; Pred. No. 1.4e-47;
                                                                                                                                               APPLICANT: Lewis, Randolph V.
APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: BW PC compatible
COMPUTER: BW PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,298
FILING DATE: 02-JUN 1995
FILING DATE: 14-MAR-1994
ATPORNEY/AGENT INFORMATION:
REGISTRATION NUMBER: 28,977
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: N. clavipes
TISSUE TYPE: minor ampullate gland
                                                                       Sequence 2, Application US/08458298
Patent No. 5756677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: 703-205-8050 INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 168; Conserv
                                                                                                                               GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                  Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: pro
HYPOTHETICAL: NO
FRAGMENT TYPE: 1nd
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                  22040-3487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY: CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-458-298-2
                                                    JS-08-458-298-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       109 AAAAGGAGQGGYGGLGSQG------TSGRGGLGGQGAGAAAAAAAAAAAGGAGQG- 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -GYGGLGSQGT-----SGRGGLG-GQGAGAAAAA-----AAAAGGAGGGYG 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   198 GLGSQGT------SGRGGLGGQGAGAAAAAAAAGGAGGGGGGG 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 832;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97; Indels
                                                                                                                                                                                                                                                                   CORTILLS AND SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release #1.0, Version #1.25 SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,747
FILING DATE: 14-MAR.1994
CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 674; DB 1;
Pred. No. 1.4e-47;
7; Mismatches 97;
                     CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SGRGGLGGQGAGAAAAAAAAAGGAGQGGYGG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
ORIGIMAL SOURCE:
ORGANISM: N. clavipes
TISSUE TYPE: minor ampullate gland
                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION:
NAME: MUTPHY Jr., Gerald M.
REGISTRATION NUMBER: 28,971
REFERENCE/DOCKET NUMBER: 1447
TELECOMMUNICATION INFORMATION:
TELEFRAN: 703-205-8000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      45.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGIH: 832 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      single
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
STRANDEDNESS: sl
                                                                                                                                                                         22040-3487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; NAME/KEY;
; LOCATION:
US-08-209-747-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         287 AA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA 673
                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

œ

```
--GAGAAAAAAAAAGGAG-----QGGYGGLGSQGTSGRGGLGGQG--AGAAAAAAAA 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----GAGQGGYGGLGSQGTSGR- 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   151 GGAGQGGYGGLGSQGTSGRGGLGGQGAG-----AAAAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11 GSMASGRGGLGGQGAGAAAAAAAAAGGAGO------GGYGGLGSQGTSG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels 147;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 738;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   246 ------GGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSG 282
                            ZIP: 514-01
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40.0%; Score 593.5; DB
41.1%; Pred. No. 4e-41;
tive 10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    204 TSGRGGLGGQGAGAAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; LOCATION: from 1 to 738; IDENTIFICATION METHOD: E (by experiment) US-08-864-038A-3
                                                                         COMPUTER: IBM Compatible
OPERATING SYSTEM: Microsoft Windows 95
SOFTWARE: WOR'D PERfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/864,038A
FILING DATE: MAY 28, 1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search completed: December 18, 2002, 16:23:03 Job time : 8.69992 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Pinctada fucata
CELL TYPE: mantle epithelial cell
                                                                                                                                                                           PRIOR APPLICATION DATA:
PRPLICATION NUMBER: JP 8-184459
FILING DATE: 15-July-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                     NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REGISTRATION NUMBER: F-56.
TELECOMMUNICATION INFORMATION:
TELEFAX: (212)953-7733
ÎNFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best_Local Similarity 41:18
Matches 167; Conservative
Mie-prefecture
                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY: peptide
                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM:
                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 3, Application US/08864038A
Patent No. 6001592
GENERAL INFORMATION:
GENERAL INFORMATION:
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
TITLE OF INVENTION: TO SAID POLYPEPTIDE
NUMBER OF SEQUENCES: 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10;
                                                                         60 GQGAGAAAAAAAAAAGGAGQ-GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAA---GA 115
             ----SGRGGLGGQGAGAAAAAAAAAGGAGQGGYG 235
                                          116 GQGGYGGLGSQGTSGRGGLGGQ-GAGAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGG 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 RGSHHHHHHGSMASGRGGLG-GQGAGAAAAAAAAAAAAAGGAGQ-GGYGGLGSQGTSGRGGLG 59
                                                                                                                                                                                                                                               Sequence 3, Application US/09249585A
Patent No. 6417002
GENERAL INFORMATION:
GENERAL INFORMATION:
TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
FILE REFERENCE: 0867/00905
CURRENT APPLICATION NUMBER: US/09/249,585A
CURRENT FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin version 3.0
SEQ ID NO 3:
LENGTH: 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       231 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGIR 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 4; Length 641;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
41.1%; Score 610.5; DB 4; Length 6.
Best Local Similarity 48.6%; Pred. No. 1.5e-42;
Matches 143; Conservative 14; Mismatches 104; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; TYPE: PRT
; ORGANISM: Epstein Barr Virus
US-09-249-585A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : 812-5 Hirano
Isshinden
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS ADDRESSE: 812-5 H11
                                                                                                                                                                                                                         RESULT 14
US-09-249-585A-3
                                                                                                                                         287 AA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 15
US-08-864-038A-3
                                                                                                                                                                         AA 673
           198
                                                                                                                                                                           672
           ò
                                        g
                                                                          ð
                                                                                                     g
                                                                                                                                         ò
                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

12;

Gaps

66

```
(without alignments)
1213.493 Million cell updates/sec
                                                                                                                     December 18, 2002, 16:21:10 ; Search time 4.01995 Seconds
                                                                                                                                                                                                   US-09-490-291-2
1485
1 MRCSHHHHHHGSMASGRGCL.....YGGLGSQGTSGIRRPAAKLN 291
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                       Title:
Perfect score:
                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                 Sequence:
                                                                                                                       Run on:
```

Total number of hits satisfying chosen parameters:

106657 seqs, 16763532 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

1: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/USOB_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/USO7_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/USO7_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/USO7_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/USOB_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/USOB_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/USOB_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:* /ptodata/1/pubpaa/US10_PUBCOMB.pep:* /ptodata/1/pubpaa/US60_NEW_PUB.pep:* /ptodata/1/pubpaa/US10_NEW_PUB.pep /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep Published_Applications_AA:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Sequence 2, Appli Sequence 3, Appli Sequence 24, Appli Sequence 1, Appli	Sequence 269, App Sequence 20, Appl Sequence 451, App Sequence 525, App Sequence 1, Appl	Sequence 16, Appl Sequence 47, Appl Sequence 3, Appl1 Sequence 106, App	sequence 52, Appl. Sequence 52, Appl. Sequence 202, Appl. Sequence 6, Appl. Sequence 74, Appl.	Sequence 581, App Sequence 146, App Sequence 614, App Sequence 614, App Sequence 614, App Sequence 312, App Sequence 312, App
10 US-09-804-898-2 10 US-09-918-951-3 9 US-10-029-217A-24 9 US-10-096-961-1	12 US-10-054-386-269 10 US-09-756-071B-20 12 US-10-052-586-451 12 US-10-052-586-525 10 US-09-735-933-1	10 US-09-823-936-16 9 US-09-975-143-47 10 US-09-826-508-3 9 US-10-066-500-106	9	12 US-10-052-586-581 10 US-09-793-306-146 9 US-09-978-697-614 9 US-09-978-697-614 12 US-10-052-586-332 10 US-09-886-055-431
4679 479 3907 2211	1300 720 1743 4440 1422	914 1200 440	440 440 1184 450 585	1435 597 520 520 520 975
22.7 22.6 22.5 22.5	21.5 20.5 20.5 20.5	19.6 18.7 18.3	18.3 18.3 18.0 17.6	17.5 16.7 16.5 16.5 16.5 16.5
337.5 336 334.5 331.5	320 319 305 304.5 299	291.5 287 277.5 272	272 272 267.5 262 261.5	259.5 247.5 245 245 245 245 245
22222	25 27 28	330	, w w w w w v 4 tv 6 t 8	3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

ALIGNMENTS

```
Sequence 4, Application US/09861597

Sequence 4. Application US/09861597

Sequence 4. Application US/09861597

Seneral Information:
APPLICANT: PHILIPPE, Michel
APPLICANT: GARSON, Jean-Pierre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE NEPERBENCE: 6388-0365-0
CURRENT FILING DATE: 2001-05-22
PRIOR PAPLICATION NUMBER: 09/247,806

PRIOR FILING DATE: 1999-02-11
PRIOR FILING DATE: 1999-02-11

PRIOR FILING DATE: 1999-02-11

SEQ ID NO 4

LENTARE: PatentIn Ver. 2.1

SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       50 QGT--SGRGGLGGQGA---GAAAAAAAAAAAAAAAAAAAAAAAAAAA 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 1036; DB 10; Length 606;
Pred. No. 6.5e-57;
4; Mismatches 23; Indels 102; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11 GSMASGRGGLGGQGAGAAAA-------AAAAAAGGAGQGGYGGLGS 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 69.8%;
Best Local Similarity 64.9%;
Matches 239; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                     ; ORGANISM: Nephila clavipes US-09-861-597-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
```

~

```
Best Local Sim
Matches 230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1, Application US/09861597

Sequence 1, Application US/09861597

Sequence 1, Sequence 2, Sequence 2, Sequence 3, Sequ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13;
                                                                                                   232
                                                                                                                                               527
                                                                                                                                                                                                271
                                                                                                                                                                                                                    467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GQGAGAAAAAAAAAA.....AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA.... 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAA-- 259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----AAAAGGAGQGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAA----- 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------AAAAAAGGAGGGKG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        46 GLGSQGTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL 96
GYGGLGSQGTSGRGGLGGQGAGAAA--------AAAAAAGGAGGGG
                                                                                               326 QGGLGGQGAGQAGAAAAAAGGAGGGYGGLGSQG-AGRGGLGGQGAGAVAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 651;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -- AAAAAGGAGQCGYGGLGSQC 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 1015; DB 10;
Pred. No. 1.3e-55;
8; Mismatches 14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11 GSMASGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      68.48;
59.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 59.2
Matches 238; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGOGAGAAAAA - - - -
                                                                                                                                                                                                                                                                                   272 YGGLGSQG 279
                                                                                                                                                                                                                                                                                                                11111111
587 YGCLGSQG 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQ ID NO 1
LENGTH: 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-861-597-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                         233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   260
                                                                                                                                      g
                                                                                                                                                                                         ò
                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                             a
  å
                                              유
                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
```

```
APPLICANT: PHILLIPPE, Michel
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Pierre
APPLICANT: ARRAUDBAU, Jean-Pierre
TITLE OF INVENTION: COSMETIC OR DERNATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
TITLE OF INVENTION: ANALOG
TITLE OF INVENTION: ANALOG
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT FILING DATE: 1999-02-11
PRIOR FILING DATE: 1999-02-11
PRIOR FILING DATE: 1999-02-11
SEQ ID NO 8
LENGTH: EACH OF SEQ ID NOS: 14
SEQ ID NO 8
LENGTH: EACH OF SEQ ID NOS: LENGTH: EACH OF SEQ ID NO 8
LENGTH: EACH OF SEQ ID NOS: LENGTH: EACH OF SEQ ID NO 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-801-597-6

Sequence 6, Application US/09861597

Sequence 6, Application US/09861597

Patent No. US2020064539A1

GENERAL INFORMATION:

APPLICANT: PHILLIPPE, Michel

APPLICANT: ARRAUDENU, Jean-Pierre

TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT

TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN

TITLE OF INVENTION: ANALOG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         137 QGAGAAAAAAAAAAGAGGGGGGGGGGGGGT-----SGRGGLGGGGGAGAAA-- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----- AAAAAAGGAGGGGYGGLGS 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98 GQGAGAAAA----------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         183 -------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            224 AAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 144;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 606;
444 GGAGGGGYGGLGNQGAGRGQQGAAAAAGGAGQGGYGGLGSQG 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 976; DB 10;
Pred. No. 2.8e-53;
4; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11 GSMASGRGGLGGQGAGAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Sequence 8, Application US/09861597
; Patent No. US20020064539A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65.7%;
59.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; ORGANISM: Nephila clavipes US-09-861-597-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
셤
```

m

```
q
                                                                                    ò
                                                                                                               g
                                                                                                                                               ö
                                                                                                                                                                             g
                                                                                                                                                                                                            à
                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                          á
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                          13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2, Application US/09861597

Patent No. US20020064539A1

GENERAL INFORMATION:

APPLICANT: PHILIPPE, Michel

APPLICANT: ARRADDRAU, Jean-Plerre

APPLICANT: ARRADDRAU, Jean-Plerre

APPLICANT: ARRADDRAU, Jean-Plerre

APPLICANT: ARRADDRAU, Jean-Plerre

TITLE OF INVENTION: COSMETIC ON DERMATOLOGICAL COMPOSITION CONTACTING AT

TITLE OF INVENTION: ANALOG

TITLE OF INVENTION: ANALOG

FILE REFERENCE: 6388-0365-0

CURRENT APPLICATION NUMBER: US/09/861,597

PRIOR FILING DATE: 1999-02-11

PRIOR FILING DATE: 1999-02-11

PRIOR FILING DATE: 1999-02-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           439 AGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AAA 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           327 GQGAAAAAAAAAGGAGGGLGSQGAGQAAAAAAAGGAGQGGYGGLGSQG----AGRGG 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QGAGAAAAAAAAAAGGAGGGYGGLGSQGT------SGRGGLGGQGAGAAAA-- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAA 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAAGGAGGGYGGLGSQGT-----SGRGGLGGQGAGAAA----- 258
                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                       11 GSMASGRGGLGGQGAGAAAA------AAAAAAGGAGQGGYGGLGS 49
                                                                                                                                                                                                                                                                                                                                                                    50 QGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLG 97
                                                                                                                                                                                                                                                                          Indels 144;
                                                                                                                                                                                                                                            Length 606;
                                                                                                                                                                                                                                                                          12;
                                                                                                                                                                                                                                           Score 970; DB 10;
Pred. No. 6.5e-53;
4; Mismatches 12;
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT FILING DATE: 2001-05-22
FRIOR APPLICATION NUMBER: 09/247,806
PRIOR FILING DATE: 1999-02-11
PRIOR FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver: 2.1
SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    551 QGAGQGAGAAAAAGGAGQGGYGGLGSQG 579
                                                                                                                                                                                                                                           65.3%;
58.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
                                                                                                                                                                           ; TYPE: PRT
; ORGANISM: Nephila clavipes
US-09-861-597-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; TYPE: PRT; ORGANISM: Nephila clavipes US-09-861-597-2
                                                                                                                                                                                                                                                        Best Local Similarity 58.9
Matches 229; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQ ID NO 2
LENGTH: 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 5
US-09-861-597-2
                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                  98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       224
                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
```

```
17;
                                            18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Plerre
APPLICANT: GARSON, Jean-Plerre
APPLICANT: GARRAUDEAU, Jean-Plerre
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: 09/247,806
PRIOR FILING DATE: 1999-02-11
PRIOR APPLICATION NUMBER: FR 98/01614
PRIOR APPLICATION NUMBER: FR 98/01614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60 GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGCLG-GQGAGAAAAAAAAAAAGGAGQG 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          174 GQGAGAAAA---------AAAAAAGGAGGGYG-GLGSQGTSGRG 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    75 GGAGQGGYG------GLGSQGTSGRG----GLGGQGAGAAAAAAAAGGAGQGGYG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                         240 GPGSAAAAAAGGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYGPGQQGPGGYGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gapa
                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10 HGSMASGRGGLG-GQ----GAGAAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG-59
                                                                                                          GLGSQGTSGRG------AAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               209 GLGGQGAGAAAAAAAAAAGGAGGAGGIGSQGTSGRG----GLGGQGAGAAAAAAAA
; Score 555.5; DB 10; Length 529;
; Pred. No. 8e-28;
11; Mismatches 90; Indels 111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OTHER INFORMATION: Description of Artificial Sequence:protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 714;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               35.3%; Score 524; DB 10; Length 71
49.2%; Pred. No. 8e-26;
ive 12; Mismatches 107; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 10, Application US/09861597 Patent No. US20020064539A1
  37.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              265 -GGAGQGGYGGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  359 PGQQGPGGYGP-GQQGPSG 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 49.2% Matches 155; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
                     Best Local Similarity 44.1
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 6
US-09-861-597-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; OTHER INFORM
US-09-861-597-10
  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                               46
                                                                                                                                                                                                                                                                                                                               122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

ä

ò g ò g ò

```
Sequence 22, Application US/09848990

Fatent No. US20020048572A1

GENERAL INFORMATION:
APPLICANT: Shan, Bei
APPLICANT: Tularik Inc.
TITLE OF INVENTION: Treatment of Hypertriglyceridemia and Other Conditions
TITLE OF INVENTION: USing Like Modulators
FILE REFERENCE: 018791-004910US
CURRENT APPLICATION NUMBER: US/09/848,990
CURRENT FILING DATE: 2001-05-03
PRIOR PELING DATE: 2000-05-03
PRIOR PELING DATE: 2000-05-03

WUMBER OF SEQ ID NOS: 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       118 GGYGGLGS------OGTSGRG--GLGGQGAGAAAAAAAAAAAGGGGGGY 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              159 GGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65 AAAAAAAAAAAG------GAGQGGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAGGAGQ 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11 GSMASGRGGLGGQGAGAAAAAAAAAGGAGQGG-----YGGLGSQGTSGRGGLGGQGAG 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OTHER INFORMATION: Description of Artificial Sequence:flexible linker
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   219 AAAAAAAG-----GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KET: MOD_RES
LOCATION: (1)..(97)
OTHER INFORMATION: Gly at positions 1-97 may be present or absent
NAME/KEY: MOD_RES
LOCATION: (105)..(201)
                                                                                                                                                                                                                                                                                                                                                                                                                                                            31.4%; Score 466.5; DB 10; Length 357; 37.8%; Pred. No. 1.5e-22; tive 14; Mismatches 137; Indels 43;
                                                  CTHER INFORMATION: MAP TO ALO78472.1

CTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL 27

CTHER INFORMATION: EXPRESSED IN LUGG, SIGNAL 39

CTHER INFORMATION: EXPRESSED IN FUT4, SIGNAL 34

CTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL 27

CTHER INFORMATION: EXPRESSED IN MULT LIVER, SIGNAL 27

CTHER INFORMATION: EXPRESSED IN HELA, SIGNAL 27

CTHER INFORMATION: EXPRESSED IN HEALY, SIGNAL 27

CTHER INFORMATION: EXPRESSED IN HEART, SIGNAL 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match
Best Local Similarity 37.8%
Matches 118; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 22
LENGTH: 201
TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 271 GYGGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       317 DGGGGGGRGFSG 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-848-990-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT PERO, Sharron G.
APPLICANT RANK, DAVIG R.
APPLICANT BRANK, SHARRON G.
APPLICANT RANK, DAVIG R.
APPLICANT REPRESENCE: RECORD RANKES BY HICROARRAY
CURRENT FILE OF INVESTION WORDER: 20,208 de4, 761
PRIOR PLINE DAVE: 2000-09-24 de4, 761
PRIOR PLINE DAVE: 2000-09-24 de4, 266
PRIOR PLINE DAVE: 2000-09-26 de4, 266
PRIOR PLINE DAVE: 2000-09-26 de4, 266
PRIOR PLINE DAVE: 2000-09-27 de6, 266
PRIOR PLINE DAVE: 2000-09-27 de6, 266
PRIOR APPLICATION NUMBER: 2000-09-27 de6, 266
PRIOR APPLICANTON NUMBER: 2001-01-30
PRIOR PLINE DAVE: 2001-01-30
PLINE DAVE: 2001-01-30
PRIOR PLINE DAVE: 2001-01-30
PRIOR PLINE DAVE: 2001-01-30
PRIOR PLINE DAVE: 2001-01-30
PRI
                                                                                                                                                                          GGYGGLGSQGTSGRG----GLGGQGAGAAAAAAAAAAGG-----AGQGGYG----- 273
                                                                                                                                                                                                                                                                                                 174 -GQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAGGAGQ 231
                  GYG-GLGSQCTSGRGGLGGQGAGAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 35807, Application US/09864761 Patent No. US20020048763A1 GENERAL INFORMATION:
APPLICANT: Penn, Sharron G. APPLICANT: Rank, David R. APPLICANT: Hanzel, David R. APPLICANT: Chen, Wensheng
                                                                                                                                                                                                                                                                                                                                                                                                                            591 GPGQGPSGPGSAAA 605
                                                                                                                                                                                                                                                                                                                                                                                   274 GLGSQGTSGIRRPAA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-864-761-35807
                                                                                                                                                                                                                                                               232
```

S

Page

```
RESULT 10
US-09-798-584-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-798-584-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-967-624-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQ ID NO 18
LENGTH: 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
                                                                                                     q
                                         g
                                                                         ŏ
                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
9
                                                                                                                                                                                                                                                                                                                                              76 GAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                        136 GQGAGAAAAAAAAAAGGAGQGCYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGG 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 135
                                                                                         Gaps
                                                                                                                                            OTHER INFORMATION: Description of Artificial Sequence:flexible linker NAME/KEY: MOD RES.
LOCATION: (1)...(97)
OTHER INFORMATION: Gly residues from positions 1-97 may be present or OTHER INFORMATION: absent
NAME/KEY: MOD RES.
LOCATION: (105)..(201)
OTHER INFORMATION: Gly residues from, positions 105-201 may be
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GRGGLGCQCACAAAAAAAAAAGCACQCGYGCLCSQCTSGRGCLGCQGAGAAAAAAAAG 75
; OTHER INFORMATION: Gly at positions 105-201 may be present or absent US-09-848-990-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 14, Application US/09760364
Sequence 14, Application US/09760364
GENERAL INFORMATION:
APPLICANT: Lehmann, Juergen Michael
APPLICANT: Lehmann, Juergen Michael
APPLICANT: Shiau, Andrew Kwan-Nan
APPLICANT: Tularik Inc.
TITLE OF INVENTION: Hypercholesterolemia
FILE REFERENCE: 018781-004110US
CURRENT APPLICATION NUMBER: US/09/760,364
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: US 60/176,398
PRIOR APPLICATION NUMBER: US 60/176,398
NUMBER OF SEQ ID NOS: 14
SSOFING DATE: 2000-01-13
LENGTH: 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29.2%; Score 433; DB 10; Length 201; llarity 42.5%; Pred. No. 1.1e-20; Conservative 0; Mismatches 91; Indels 56
                                                     29.2%; Score 433; DB 10; Length 201; 42.5%; Pred. No. 1.1e-20; Live 0; Mismatches 91; Indels 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                         256 AAAAAAAAGGAGGGYGG 274
                                                                                                                                                                                                                                                                                                                                                                                                                       183 GGGGGGGGGGGGGG 201
                                                                      Best Local Similarity 42.5
Matches 110; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Sim.
Matches 110;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-760-364-14
                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9/
                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                      a
                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                      δ
```

```
Sequence 18, Application US/09798584

Patent No. US20020102676A1

GENERAL INFORMATION:

APPLICANT: Mu, David

APPLICANT: Tularik Inc.

APPLICANT: Tularik Inc.

APPLICANT: Tularik Inc.

APPLICANT: Tularik Inc.

TITLE OF INVENTION: KCNB: A No. US20020102676Alel Potassium Channel Protein

FILE REFERENCE: 018781-004010US

CURRENT APPLICATION NUMBER: US/09/798,584

CURRENT FILING DATE: 2001-03-03

NUMBER OF SEQ ID NOS: 18

SOFTWARE: Patentin Ver. 2.1
                                                                                                                       76 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      136 GOGAGAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGG 195
                     OTHER INFORMATION: Description of Artificial Sequence:poly-Gly linker
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16 GRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAA 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY: MOD_RES
LOCATION: (6)..(200)
OTHER INFORMATION: Gly at positions 6-200 may be present or absent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29.1%; Score 431.5; DB 10; Length 200; 42.5%; Pred. No. 1.3e-20; tive 0; Mismatches 90; Indels 59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 19, Application US/09967624
Patent No. US20020142325A1
GENERAL INFORMATION:
APPLICANT: Liao, X. Charlene
APPLICANT: Chu, Peter
APPLICANT: Pardo, Jorge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 GGGGGGGGGGGGGG 200
                                                                                                                                                                                         256 AAAAAAAAGGAGQGGYGG 274
                                                                                                                                                                                                                                     183 GGGGGGGGGGGGGG 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           156 AAAAAAAAGGAGGGGYGG 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 42.5
Matches 110; Conservative
```

ဖ

```
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
US-09-861-597-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQ ID NO 7
LENGIH: 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           윱
                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           76 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 136 GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGG 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   196 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGA 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: MOD_RES
LOCATION: (6)..(200)
OTHER INFORMATION: Gly residues from position 6 to 200 may be present
OTHER INFORMATION: or absent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16 GRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAG 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 18, Application US/09998667

Batent No. US2002014674741

GENERAL INFORMATION:

APPLICANT: Masuda Esteban

APPLICANT: Liao, X. Charlene

APPLICANT: Liao, X. Charlene

APPLICANT: Chu, Peter

APPLICANT: Rigel Pharmaceuticals, Incorporated

APPLICANT: Rigel Pharmaceuticals, Incorporated

APPLICANT: Rigel Pharmaceuticals, Incorporated

APPLICANT: Pardo, Jorge

APPLICANT: Pardo, Jorge

APPLICANT: Now INVERSIT WOUGHOUS

FILE REFERENCE: 021044-000600US

FILE REFERENCE: 021044-00600US

CURRENT FILING DATE: 2001-12-03

PRIOR PILING DATE: 2001-12-03

PRIOR FILING DATE: 2001-04-06

NUMBER OF SEC ID NOS: 18

SCOFTWARE: PATENTIN ONS: 18

LENGTH: 200

TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                             OTHER INFORMATION: Description of Artificial Sequence:poly Gly tag
OTHER INFORMATION: flexible linker
                              APPLICANT: Wu, Jun
APPLICANT: Rigel Pharmaceuticals, Inc.
APPLICANT: Rigel Pharmaceuticals, Inc.
TITLE OF INVENTION: PAK2: modulators of Lymphocyte Activation
FILE REFERENCE: 021044-00070000S
CURRENT APPLICATION NUMBER: US/09/967,624
CURRENT FILING DATE: 2001-09-28
PRIOR FILING DATE: 2001-03-30
NUMBER OF SEQ ID NOS: 19
SEQ ID NOS: 19
SEQ ID NO 19
LENGTH: 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 29.1%; Score 431.5; DB 10; Length 200; Best Local Similarity 42.5%; Pred. No. 1.3e-20; Matches 110; Conservative 0; Mismatches 90; Indels 59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182 GGGGGGGGGGGGGG 200
                                                                                                                                                                                                                                                                                                                     ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     256 AAAAAAAAGGAGGGGGG 274
Li, Congfen
Zhao, Haoran
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-998-667-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-967-624-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

```
ġ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: ARRANDEM: Jean-Claude
APPLICANT: ARRANDEM: Jean-Plarre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REPERENCE: 6388-0385-0
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT FILLING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: 09/247,806
PRIOR FILLING DATE: 1999-02-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               76 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 196 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGA 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  152 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   212 GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGG 271
                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURE: OTHER INFORMATION: Description of Artificial Sequence:flexible linker
                                                                                                                                                                                                                                                                                           : LOCATION: (6)..(200)

: OTHER INFORMATION: Gly at positions 6-200 may be present or absent

US-09-998-667-18
                                                                                                                                                                                                 Length 200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 101;
                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                         DB 10;
                                                                                                                                                                                         Query Match 29.1%; Score 431.5; DB 10;
Best Local Similarity 42.5%; Pred. No. 1.3e-20;
Matches 110; Conservative 0; Mismatches 90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 28.4%; Score 421; DB 10; Best Local Similarity 74.0%; Pred. No. 3.5e-20; Matches 94; Conservative 1; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION NUMBER: FR 98/01614
PRIOR FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTIN VEr. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 7, Application US/09861597
Patent No. US20020064539A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                256 AAAAAAAAGGAGGGGGG 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182 GGGGGGGGGGGGGG 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; ORGANISM: Nephila clavipes US-09-861-597-7
                                             FEATURE:
NAME/KEY: MOD_RES
LOCATION: (6)..(20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                272 YGGLGSQ 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
```

```
PPRIOR RELIGIOR PRINCIPAL RELIGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR
PRIOR
PRIOR
                                                                                                                                                                                             GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: PHILLIPPE, Michel
APPLICANT: PRILLIPPE, Michel
APPLICANT: ARRADDEAU, Jean-Pierre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT PILLING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: PR 98/01614
PRIOR FILING DATE: 1999-02-11
PRIOR PILLING DATE: 1998-02-11
SPIOR APPLICATION NUMBER: FR 98/01614
SOFTWARE: Patentin Ver. 2.1
SOFTWARE: Patentin Ver. 2.1
LENGTH: 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430RL1
CURRENT APPLICATION NUMBER: US/10/052,586
CURRENT FILING DATE: 2002-01-15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             212 GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGG 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 27.7%; Score 412; DB 10; Length 101; Best Local Similarity 73.6%; Pred. No. 1.2e-19; Matches 92; Conservative 1; Mismatches 6; Indels 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
0S-10-052-586-97
Sequence 97, Application US/10052586
Patent No. US20020127584A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
                                                                                                                                                     Sequence 5, Application US/09861597
Patent No. US20020064539A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pan, James
Smith, Victoria
Watanabe, Colin K.
Wood, William I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Goddard, Audrey
Godowski, Paul J.
Gurney, Austin L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; TYPE: PRT
; ORGANISM: Nephila clavipes
US-09-861-597-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Desnoyers, Luc
95 YGGLGSQ 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   272 YGGLG 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97 YGGLG 101
                                                                                                     RESULT 14
US-09-861-597-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
PRIOR FILING DATE: 1997-10-17
PRIOR PILING DATE: 1997-10-24
PRIOR PILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR PILING DATE: 1997-10-24
PRIOR PILING DATE: 1997-10-28
PRIOR PILING DATE: 1997-10-31
PRIOR PILING DATE: 1997-11-34
PRIOR PILING DATE: 1997-11-37
PRIOR PILING DATE: 1997-12-11
PRIOR PILING DATE: 1996-03-10
PRIOR PILING DATE: 1996-03-11
PRIOR PILING DATE: 1996-04-08
```

```
PRIOR APPLICATION NUMBER: 60/082797

REIOR APPLICATION NUMBER: 60/08332

REIOR FILING DATE: 1998-04-29

PRIOR FILING DATE: 1998-04-29

PRIOR PELICATION NUMBER: 60/08332

PRIOR FILING DATE: 1998-04-29

REIOR APPLICATION NUMBER: 60/083495

PRIOR FILING DATE: 1998-04-29

REIOR APPLICATION NUMBER: 60/083495

PRIOR FILING DATE: 1998-04-29

REIOR PLING DATE: 1998-04-29

REIOR FILING DATE: 1998-05-06

PRIOR FILING DATE: 1998-05-07

PRIOR FILING DATE: 1998-05-07

PRIOR FILING DATE: 1998-05-07

PRIOR FILING DATE: 1998-05-07

PRIOR PLING DATE: 1998-05-15

PRIOR PLING DATE: 1998-05-16

PRIOR PLING DATE: 1998-05-18

PRIOR PLING DATE: 1998-05-07

PRIOR PLING DATE: 1998-05-07

PRIOR PLING DATE: 1998-05-07

PRIOR PLING DATE: 1998-05-07

PRIOR PLING DATE: 1998-05-05

PRIOR PLING DATE: 1998-0
```

1108 GAAGAAATCAGTCAGCCGCAGTGAAGAGCTCAGAAAAGAAGCAAGAACAATTAAAAACGG-G 1166 132 G-----GLGGQG---AGAAAAAAAAAAAGGA----GQGGYGGLG 162 106 ------AAAAAAGGA-------131 928 AGTIGIAGAAAGIGAAAAAGGIGAIGCACCAGAITIAGIIGAIGAIGGAGAAGAIGAAA 987 12; Mismatches 122; Indels 113; Gaps 163 SQGTSGRGGLGGQGAGAAAAAAAAAAAGGGGGGGGGCTGC-----SQGTSGRGGLGGQGA 216 GAAAAAAAAAAAGGAG-----QGGYGGLGSQGTSGRGG-----LGGQGAGAAAAAAAA Length 1894; DB 12; Pred. No. 1.8e-18 27.4%; Score 407.5; 35.0%; Pred. No. 1.8 Search completed: December 18, 2002, 16:38:50 Job time: 8.01995 secs R FILING DATE: 1998-06-10

R APPLICATION NUMBER: 60/088825

R FILING DATE: 1998-06-10

R APPLICATION NUMBER: 60/088826

R FILING DATE: 1998-06-10

R APPLICATION NUMBER: 60/088861

R FILING DATE: 1998-06-11 15 SGRGGLGGQGAGAAAAAAAAAGGAG----APPLICATION NUMBER: 60/088811 FILING DATE: 1998-06-10 APPLICATION NUMBER: 60/088824 FILING DATE: 1998-06-11 APPLICATION NUMBER: 60/088876 FILING DATE: 1998-06-11
APPLICATION NUMBER: 60/089090
PILING DATE: 1998-06-12
APPLICATION NUMBER: 60/089105 FILING DATE: 1998-06-12
APPLICATION NUMBER: 60/089512
FILING DATE: 1998-06-16
APPLICATION NUMBER: 60/089514 PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089908 APPLICATION NUMBER: 60/088863 APPLICATION NUMBER: 60/089538 FILING DATE: 1998-06-17 APPLICATION NUMBER: 60/089598 1287 AGTATGAAGCTTTGAGGAAG 1306 264 AGGA-GQGGYGGLGSQGTSG 282 FILING DATE: 1998-06-16 Best Local Similarity 35.0% Matches 133; Conservative Query Match PRIOR
PRIOR à 셤 ò g ò ద ò g ò g ò

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
```

OM protein - protein search, using sw model

December 18, 2002, 16:13:49; Search time 8.59824 Seconds (without alignments) 3253.588 Million cell updates/sec Run on:

US-09-490-291-2 1485 1 MRGSHHHHHHGSMASGRGGL.....XGGLGSQGTSGIRRPAAKLN 291 Title: Perfect score: Sequence:

Scoring table:

BLOSUM62 Gapop 10.0 , Gapext 0.5

283224 seqs, 96134422 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	. =	fibroin - Chinese	nuclear antigen EB	hypothetical glyci	cal		hypothetical glyci			hypothetical glyci		qîycine-rich ceîl	hypothetical glyci	_	-	_	_	_	_	_	_	_	c		_	_	_	hypothetical glyci	4
SOFFICE	OI (I	A36068	T31328	QQBE31	F70806	D70807	A44112	E70806	B70807	H70846	A70812	A70869	A26099	A70896	F70963	A70934	E70895	G70846	E70820	D70916	C70974	A70807	B70812	S01820	A70770	A70900	E70917	A70893	H70987	KNMU
	DB	7	~	Н	7	7	~1	7	~	~	7	7	-	~	7	~	7	7	7	~	~	~	7	-	7	~	~	7	~	Н
	Query Match Length	71	2639	641	1901	1489	627	1381	1079	1538	749	1660	384	853	778	1306	167	484	923	491	731	714	882.	465	603	576	1329	299	914	338
dР	Query	٠.		4	37.5	37.4	37.3	9	36.6	36.2	36.0	35.4	'n.	35.0	4	4	4	4.	4		m.	ë.	ë.	m.	33.4	•		m	ش	33.0
	Score	1015	617.5	610.5	556.5	556	554	544.5	543.5	538	535	525.5	525	519.5	516.5	514	512.5	512	507.5	503	503	499	498.5	9	496	495	495	494	464	490
	Result No.	ч	7	Ю	4	S	9	7	89	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	. 59

RESULT 2 T31328

0.2	glycine-rich prote hypothetical glyci hypothetical glyci		hypothetical glyci hypothetical glyci hypothetical glyci	hypothetical glycl hypothetical glycl hypothetical glycl	
F70824 S31415	T49109 A70989 C70720	F70868 H70839 G70917	F70620 H70589 B70523	F70580 E70824 G70545	H70816 E70835
7 7	900	000	000	000	00
801 291	396 618 498	694 496 741	1011 615 591	532 783 594	606
33.0	33.0 33.0	32.8 32.8 32.8	32.7 32.6 32.4	32.4	32.2
489.5	489.5 489.5 489	487.5 487 486.5	486 484 481	480.5 480.5 479	477.5
30 31	3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	35 37	38 40 40	4 4 4 1 2 2 1 3 2 5	444

ALIGNMENTS

RESULT 1 A36068 major ampullate fibroin protein - orb spider (Nephila clavipes) (fragment) C; Species: Nephila clavipes C; Species: Nephila clavipes C; Accession: A36068 E; Xi, W. M.; Lewis, R.V. Broc. Natl. Acad. S.V. A; Title: Structure of a protein superfiber: spider dragline silk. A; Reference number: A36068; MUID:90384959; PMID:2402494 A; Reference number: A36068; MUID:90384959; PMID:2402494 A; Scatus: preliminary A; Mccession: A36068 A; Status: preliminary A; Mcsiducs: 1-718 < XUA> A; Residuces: 1-718 < XUA> A; Note: the authors translated the codon GGT for residue 292 as Gln, GTA for residue	Query Match 68.4%; Score 1015; DB 2; Length 718; Best Local Similarity 59.2%; Pred. No. 1.9e-53; Matches 238; Conservative 8; Mismatches 14; Indels 142; Gaps 13;	11 GSMASGRGGLGGQGAGAAAA	GLGSQCTSGRGGLGGQGAGAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGL 96 :	GCQGAGAAAAAA	GOGAGAAAAAAAA AGGAGQGGGGGGGGGGGGGGGG		GCYGGLGSQGAGRGGGAGAGGGGGGGGGGGGGGGGGGGGAGAAAA- 259	260
LT 1 68 60 ecies te: 0 cession M.; M.; M.; M.; M.; M.; M.; M.; M.; Sidue ecession feren feren feren feren feren sidue feren sidue sid sidue sidue sid sidue sidue sid sidue sid sidue sidue sidue sidue sidue sidu	ery M st Lo tches	11 93	46	97	136 267	183	221 385	260
RESULT 1 A36068 major am C; Specie C; Date: C; Access R; Xu, M. Proc. N A; Title: A; Refere A; Refere A; Refere A; Resultus	On Be	Qy Dp	oy D	Oy Dp	oy Op	oy B	Qy Db	Oy Dp

10;

```
Pypothetical glycine-rich protein Rv3508 - Mycobacterium tuberculosis (strain H37RV) (Species: Mycobacterium tuberculosis (Species: Mycobacterium tuberculosis (Species: Mycobacterium tuberculosis (Spacesion: F70806 (Spaces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Residues: 1-1901 <COL>
A; Residues: 1-1901 <COL>
A; Cross-references: GB:AL022022; GB:AL123456; NID:g3261554; PIDN:CAA17745.1; PID:g292
A; Experimental source: strain H37Rv
C; Genetics:
                 R;Sample, J.; Hummel, M.; Braun, D.; Birkenbach, M.; Kieff, E.
Proc. Natl. Acad. Sci. U.53... 83, 5096-5100, 1986
A;Title: Nucleotide sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: A;Reference number: S42440; MUID:86259739; PMID:3460083
                                                                                                                                               A; Accession: S42440
A; Status: preliminary
A; Molecule type: mRNA
A; Residues: 1-66 < SARNA
A; Residues: 1-66 < SABL: M13941; NID: 9330399; PIDN: AAA45889.1; PID: 9555157
C; Superfamily: Epstein-Barr virus nuclear antigen
C; Keywords: DNA binding; transcription regulation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      175 QGAGAAAAAAAAAAGGAGGGYGGLGSQGTSG----RGGLGGQGAGAAAAAAAAAAAGGAG 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           120 YGGL----GSQGTSGRGGLGQQAGAAAAAAAAGGAGO-----GGYGGLGSQGTSGR 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 RGSHHHHHHGSMASGRGGLG-GQGAGAAAAAAAAGGAGQ-GGYGGLGSQGTSGRGGLG 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------GGLGGQGAG 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65 AAAAAAAAAAGAGGGY----GGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAG-QGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   476 AGAGGAGGAAGTGGTGGVVGAAGKAGIGGTGGQGGAGGAGSAGTDATATGATGGTGFSGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            116 GQGGYGGLGSQGTSGRGGLGGQ-GAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              231 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGIR 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1901;
                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 641;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 41.1%; Score 610.5; DB 1; Best Local Similarity 48.6%; Pred. No. 1.2e-29; Matches 143; Conservative 14; Mismatches 104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 37.5%; Score 556.5; DB 2; Best Local Similarity 44.9%; Pred. No. 4e-26; Matches 140; Conservative 13; Mismatches 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14 ASGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGR--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gene: Rv3508
;Superfamily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C. Species: human herpesvirus 4
C. Species: human herpesvirus 4
C. Species: human herpesvirus 4, Epstein-Barr virus
C. Species: human herpesvirus 4, Epstein-Barr virus
C. Species: human herpesvirus 4, Epstein-Barr virus
C. Species: buman herpesvirus 4, Epstein-Barr virus
C. Species: Seducies: S42440; A03773; S33021
R. Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 2145, 1983
A; Title: Sequence analysis of the 17,166 bp ECORI fragment C of B95-8 Epstein-Barr virus
A; Reference number: A93065; MUID:85035713; PMID:6092825
A; Residues: 1-641 < BAN>
A; Residues: 1-641 < CAN>
A; Reference: Strain B95-8
A; Crost and expression of the B95-8 Epstein-Barr virus genome.
A; Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A; Reference number: A03794; MUID:84270667; PMID:6087149
A; Contents: annotation; protein coding region
                                                                                                                                                                                                                       Anth
flbroin - Chinese oak silkmoth
C;Species: Attheraea pernyl (Chinese oak silkmoth)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T31328
R;Scattsu, H.; Tamura, T.; Yukuhiro, K.
R;Scattsu, Presiminary, Translated from GB/EMBL/DDBJ
R;Reference number: 220995
R;Scatus: preliminary; translated from GB/EMBL/DDBJ
R;Reference type: DNA
R;References: EMBL:AF083334; NID:g3450882; PID:g3450883; PIDN:AAC32606.1
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1382 SAAAAAAAAAAAAAASCAGGSGGYGGYGSDSAAAAAAAAAAAAAAGSGAGGVGGGYGWGDG 1441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -------GLGSQGTSGRGGLGG------ 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------SGRGGLGG------QGAGAAAAAAA---AGGAGQGGYG- 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        213 QGAGAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAG-GAGQGG 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---QGAGAAAAAAAAAGGAGQGGYGGLGSQGT----- 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          53 ----SGRGGLGG------QGAGAAAAAAAAAGGGGGYGGLGSQGT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----QGAGAAAAAAAAGAGGGYGGLGS--------QGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           137 --QGAGAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     90; Indels 163;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 2639;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             41.6%; Score 617.5; DB 2
39.8%; Pred. No. 1.3e-29;
iive 10; Mismatches 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    272 YGGLGSQGTSGIRRPAA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best_Local Similarity 39.89
Matches 174; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14 ASGRGGLGG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Introns: 14/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

m

```
C; Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change '10-Mar-1998 C; Accession: A44112; S27824 R. Hinman, M. B.; Lewis, R.V.
J. Biol. Chem. 267, 19320-19324, 1992
A; Title: Isolation of a clone encoding a second dragline silk fibroin. Nephila clavip A; Reference number: A44112; MUID:92406876; PMID:1527052
A; Status: preliminary
A; Molecule type: mRNA
A; Residues: 1-627 < RINNA
A; Rote: sequence extracted from NCBI backbone (NCBIP:113893)
B; Hinman, M. B.; Lewis, R.V.
A; Note: sequence extracted from NCBI backbone (NCBIP:113893)
B; Hinman, M. B.; Lewis, R.V.
Submitted to the EMBL Data Library, May 1992
A; Reference number: S27824
A; Reference number: S27824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                D.; Gordon
.; Holroyd,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Molecule type: DNA
A;Reaidues: 1-1381 <COL>
A;Cross-references: GB:ALO22022; GB:ALJ23456; NID:93261554; PIDN:CAA17744.1; PID:9292
A;Experimental source: strain H37Rv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hypothetical glycine-rich protein Rv3507 - Mycobacterium tuberculosis (strain H37RV) C;Species: Mycobacterium tuberculosis C;Species: I7-Jul-1998 #text_change.20-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C. Accession: E70806
R.Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Go; Connor, R.; Brosch, R.; Pevlin, K.; Feltwell, T.; Gentles, S.; Hamiin, N.; Holri Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Matther 393, 537-544, 1998
A.Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete A. Reference number: A70500; MuID:98295987; PMID:9634230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Status: preliminary; nucleic acid sequence not shown; translation not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51 GTSGRG----GLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQCTSGRG----GLGGQGAG 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      103 AAAAAAAAAAGGAGGGYGGLGSQGTSGRG----GLGGQGAGAAAA---------- 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    206 AAAAAAASGPGQQGPGGYGP-GQQGPGGYGPGQQGLSGPGSAAAAAAAGPGQQGPGGYGP 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----AAAAAAGGGGGYG-GLGSQGTSGRGLGGQGAGAAAAAAAAAGGAGQ 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------GLGGGGAG----GAGQGGYGG 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               383 SAAAAAAAAAGGGYGPGGYAPGQQGPSGPSSASAAAAAAAAAAGPGGYGPGGGPGG 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 GSHHHHHHGSMASGRGGLGGQGAGAAAAAAAAAAAGGAGQ---GGYG------GLGSQ 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15; Mismatches 101; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 37.3%; Score 554; DB 2; Best Local Similarity 44.1%; Pred. No. 2.5e-26; Matches 166; Conservative 15; Mismatches 101.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Cross-references: EMBL: M92913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                443 YAPGQQGPSGPGSAAA 458
                                                                                                                                                                                                                                                                                                                                                                                                                                      Molecule type: mRNA; Residues: 19-627 <HI2>
                                                                                                                                                                                                                                                                                                                                                                                                                 Accession: S27824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Procession: State Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Mature 393, 357-544, 1998
Nature 393, 357-544, 1998
N.Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A.Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A;Reference number: A70500; MUID:98295987; PMID:9634230
A.Accession: D70807
A.Accession: D70807
A.Status: preliminary; nucleic acid sequence not shown; translation not shown A;Residues: 1-1489 CCOL>
A.Coosa-references: GB:AL022022; GB:AL123456; NID:93261554; PIDN:CAA17751.1; PID:9292445
C.Genetics:
C.Genetics: Ry3514
C.; Superfamily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                           hypothetical glycine-rich protein Rv3514 - Mycobacterium tuberculosis (strain H37RV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          spidroin 2, dragline silk fibroin - orb spider (Nephila clavipes) (fragment)
N;Alternate names: silk fibroin, dragline
C;Species: Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: D70807
                                                                                                                                                                                                           653 GAGGAIGTGGTGGAVGSVGNAGIGGTGGTGGVGGAGGAGAAAAAGSSATGGAGFAGGAGG 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --- GGYGGLGSQGTSGRGGL 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----TSGR 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---GAGAA 66
     94 GGLGGQGAGAAAAAAAAAAGAGQGGYGG-LGSQGTSGRGGLGGQ----GAGAAAAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGLGGQG--AGAAAAAAAAAAGG-AGQGGYGGLGSQGT---SGRGGLGGQGAGAAAAAA
                                                                                                                                                                              224 AAAGGAGQGGYGG-----LGSQGTSGRGGLGGQ-GAGAAAAAAAAAAAGGA----GQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         637 PGATGGTGFAGGAGGAGKAGGSSSAGGTNSSGSAGGTGRQSGTGGAGGAGADNPTGIGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 232 GGYGGLGSQGT---SGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG-LGSQGTSGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 1489;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    37.4%; Score 556; DB 2; Length 14
39.3%; Pred. No. 3.5e-26;
tive 20; Mismatches 109; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15 SCRGGLGGQGAGAAAAAAAAAGGAGQGGYGG-LGSQGTSGRGGLGGQ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A-------AAAAAAGGA-------GQGGYGGLGSQG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                713 EGGAG--GNSGV 722
                                                                                                                                                                                                                                                                                 272 YGGLGSQGTSGI 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Sim
Matches 140;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                757
                                                                                                                          g
                                                                                                                                                                              ö
                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
```

```
hypothetical glycine-rich protein Rv3345c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Ju1-1998 #sequence_revision 17-Ju1-1998 #text_change 20-Jun-2000
C;Accession: H70846
R;Cole, S.T.: Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon; Connor, R.; Davies, R.; Devihin, K.; Feltwell, T.; Gentles, S.; Hanris, D.; Gordon; Connor, R.; Davies, T.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 333, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Residues: 1-1538 <COL>
A;Coss-references: GB:AL021841; GB:AL123456; NID:93261517; PIDN:CAA17117.1; PID:9326
C;Genetios:
A;Genetios:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hypothetical glycine-rich protein Rv0833 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1150 ETGFCGAGGNGGYGGPG--GPEGNGGLGGNGGAGGNGGVSTTGGDGGAGGKGGNGGDGN 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1090 VGGVGGAGGDGGAGGVGGFGGQGGIGGEGRTGGNGGSGGDGGGISLGGNGGLGGNGGVS 1149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQ--- 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11 GSMASGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 70
                                                                                                                                                                                                                                   AAGGAG-----QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGL--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       71 AAAAGGA----GQGGYGGLGSQG-----TSGRGGLGGQGAGAAAAAAAAAGG----
                                                                                                                                                                                                    -------GSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----GTSGRGGLGGQGAGAAAAAAAAGG------AGQGGYGGLGSQGTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      206 GRGGLGGQGAGAAAAAA---AAAGGAGQGGYGGLGSQGTS-GRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                247 GLGGQG-------AGAAAAAAAAAAAAGGA-GQGGY-GGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                               Length 1538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 36.2%; Score 538; DB 2; L/Best Local Similarity 38.9%; Pred. No. 4.2e-25; Matches 126; Conservative 13; Mismatches 131;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VVNVTAGHGGNGGNGGNGSAG 1351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               262 --- AAAGGAGQGGYGGLGSQGTSG 282
                                                                                           149
                                                                                                                                                                                                    200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            165
                                                                                        ð
                                                                                                                                           g
                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A Status: preliminary; nucleic acid sequence not shown; translation not shown A; Molecule type: DNA A; Molecule type: DNA A; Molecule type: DNA A; Molecule type: DNA A; Molecule type: 1-1079 <COL>
A; Residues: 1-1079 <COL>
A; Experimental source: Strain H37Rv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hypothetical glycine-rich protein Rv3512 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: B70807
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, R.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genom A;Accession: B70807
                                                                                                                                                                       13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14;
                                                                                                                                                                                                                                                                                                                                                                   140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 222 AAAA----AGGAGQGGYGGLGSQGTSGRGGLGGQ------------GAGAAAAAA 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                141 AAAAAAAAAAGGGG--YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAG----GAGQG 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GYGGLGSQ------GTSGRGGLGGQ--------GAGAAAA 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       57 GLGGGGAGAAAAAAAAAAGGAG-QGGYGGLGSQ------GTSGRGGLGGQG- 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             277 GIGGNGGNGAIGGAGGDAGGSGNSGGNGGIGGKGGNAGAGGAAGSNGGIVGANGIGGDGG 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -AGAAAAAAAAAGGAGGGYGGLGSQCTSGRGGLGGQG------AGAAAAAAAAAAAAAAAA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                               1 MRGSHHHHHHGSMASGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGL--GSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 36.7%; Score 544.5; DB 2; Length 1381; Best Local Similarity 37.8%; Pred. No. 1.6e-25; Matches 144; Conservative 17; Mismatches 115; Indels 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             :68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 1079;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36.6%; Score 543.5; DB 2;
40.7%; Pred. No. 1.5e-25;
ive 12; Mismatches 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14 ASGRGGLGGQGA-GAAAAAAAAAGGAGQGGYGGLGSQGTS---
                                                C; Superfamily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         262 AAAGGAGQGGYGGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 40.79
Matches 144; Conservative
                            A; Gene: Rv3507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :Gene: Rv3512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      101
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
```

ó

```
RESULT 13
A70896
A70896
C. Special glycine-rich protein Rv1091 - Mycobacterium tuberculosis (strain H37RV)
C. Species: Mycobacterium tuberculosis
C. Species: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 24-Nov-1999
C. Accession: A70896
R. Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                glycine-rich cell wall structural protein - garden petunia
C:Species: Petunia x hybrida (garden petunia)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A26099
R:Condit, C.M.; Meagher, R.B.
Nature 323, 178-181, 1986
A:Title: A gene encoding a novel glycine-rich structural protein of petunia.
A:Reference number: A26099
                                                                                                                                                                                               GGGDGGAGGVGGHGGDGGVGGAAPSGTVGSHGTGGVGGDGGLG-GAGGVGGAGGNGGIGI 1140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1141 TVGGAGGAGGNGGDPGAGGRGGLGGDSGNGTSAANGVDASKHGPLTGGDGGVGGNGAKAA 1200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1201. AAGGDGGQGDG--GNAGLFGDGGAGGDGADGTAAEALGGDGGAGGAGGKGGDAGDIGDG 1258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GTSGRGGLGGQGA------GAAAAAAAAAAAAAGGA--GQGGYGGLGSQGTSGRGGL- 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GAAAAAAAAAAAAAAGGAGGGYGG-----LGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAAGG--- 114
                                                                                                                                                                                                                                                                               --AGQGGYGGLGSQ-GTSGRGGLGGQ-------GAGAGGYGGLGSQ-GAGAAAAA 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             238 GGGVGGGVGGGAAGGGGGGGGGGGGGGGLGGGSGHGGGFGAGGGVGGGAAGGVGGGGGF 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMASGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAA 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ASGRGGLGGQ-GAGAAAAAAAAAAGGAGGGYGGLGSQGTSGR-----GGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                        AAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAAGGA----LGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          120 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGGGGGGGGGGGS----QGTSGRGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 384;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 525, DB 1; Length 38
Pred. No. 9e-25;
2; Mismatches 136; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Residues: 1-384 <CON>
C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         232 GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLG 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGGGGGGVGGGSGHGGGFGAGGGVGGAGGGLGGGGGGAGGGGGIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1319 GGGGGVGGHGGAGGDAGMNGGGGTGGQGGNGAAGGAG 1355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    249 --- GGQGAGAAAAAAAAGGAGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     35.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Sim
Matches 125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12
                                                                                                                                                     64
                                                                                                                                                                                                                  1082
                                                                                                                                                                                                                                                                                                                                                                                                     148
                                   14
                                                                                                                                                                                                                                                                               115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   298
                                                                                                                                                                                                            윱
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                   ò
                                                                                      g
                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ayobecherical glycine-rich protein Rv2490c - Mycobacterium tuberculosis (strain H37RV)
C; Species: Mycobacterium tuberculosis
R; Spates, R; Parkhill, J; Garnier, T; Churcher, C; Harris, D.; Gordon, S; Connor, R; Davies, R; Pevilin, K; Feltwell, T; Gentles, S; Hamlin, N; Holroyd, S. Rature 393, 537-544, 1998
A; Authors: Sqares, R; Sulston, JE; Taylor, K; Whitehead, S; Barrell, B.G.
A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A; Reference number: A7060; MuID:98295987; PMID:9634230
A; Accession: A7066
A; Residues: 1-1660 <COL>
A; Residues: 1-1660 <COL>
A; Residues: 1-1660 <COL>
A; Resperimental source: Strain H37Rv
C; Genetics: Strain H37Rv
            Ricole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Ralandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A.Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A.Title: Deciphering the blology of Mycobacterium tuberculosis from the complete genome A.Title: Deciphering the blology of Mycobacterium tuberculosis from the complete genome A.Accession: A70802
A.Reference number: A70500; MUID:98295987; PMID:9634230
A.Molecule type: DNA
A.Molecule type: DNA
A.Molecule type: DNA
A.Residues: 1-749 <CGL>
A.CCL>
A.CCOSS-references: GB.AL022004; GB.AL123456; NID:g3261550; PIDN:CAA17639.1; PID:g291689
A.Experimental source: strain H37Rv
C.; Genetics:
A.Genetics:
A.Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         56 GCLGGQGAGAAAAAAAAAAAGGA-----GQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      109 AAAAGGAGGGYGGLGSQG-----TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGL 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122 GGTGGGAGSGGTGGVGGTGGAGGLFASGGAGGAGGSGGTGGAGGTGGAGGLFGAGGAGGL 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GSQG--TSGRGGLGGQ-----GAGAAAAAAAAAAGGA-GQGGYGGL--GSQGTSGR 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  242 GGAAGTFGDTGNSGGAGGAGGKAGLLFGSGGAGGSGGAGGFANGSTGGAGGAGGAGLIG 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16 GRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQ------GTSGR 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 749;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36.0%; Score 535; DB 2; Length 74 39.9%; Pred. No. 3.7e-25; Indels ive 9; Mismatches 123; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 35.4%; Score 525.5; DB 2; Best Local Similarity 40.7%; Pred. No. 2.4e-24; Matches 137; Conservative 12; Mismatches 117;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Superfamily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity Surversetive Matches 133; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
Accession: A70812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rv2490c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

., 5

237

265

```
AyObtetical glycine-rich protein Rv0578c - Mycobacterium tuberculosis (strain H37RV) C; Species: Mycobacterium tuberculosis C; Species: Nycobacterium tuberculosis C; Species: Nycobacterium tuberculosis C; Species: Nycobacterium tuberculosis C; Rarris, Dovies, Nycobacterium tuberculosis C; Harris, D; Gordon; C; Davies, R; Devlin, K; Feltwell, T; Gentles, S; Hamilin, N; Holroyd, Rajandream, M.A.: Rogers, J; Rutter, S; Seeger, K; Skelton, Ş; Squares, S. Nature 393, 537-544, 1998
A Attitle: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A; Reference number: A70500; MUID: 98295987; PMID: 9634230
A; Reference number: A70500
A; Redecession: A70934
A; Scatus: preliminary; nucleic acid sequence not shown; translation not shown A; Molecule type: DNA
A; Redeces: GB:AL021942; GB:AL123456; NID: 93242298; PIDN: CAA17449.1; PID: 9290
A; Experimental source: strain H37RV
C; Superfamily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    51 GTSGRGGLGGQGAGAAAAAAAAAGGAGGGGYGGLGSQGTS-----GRGGLGGQG-- 100
                      154 GQGGYGGL-GSQGTSGRGGLGGQGA-GAAAAAAAAAAAGAGQGGYGGLGSQGTSGRGCLG
                                                                                                                                 574 GGAGGEGLTDGAGTAEGGTGGLGGLGGVGGTGGMGGSGGVGGNGGAAGSLIGLGGGGGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         905 GDHALSGNGA-AGGNGGNGSLRGSGGAGGHGGNGGNASRGMGGDGGTGGAGGNAGQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  110 AAAGGAG-----QGGYGGLGS-QGTSGRGGLGGQGA----GAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                3 GSHHHHHHGSMASGRGGLGGQGA--GAAAAAAAAAGG-----AGQGGYGGLGSQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 101 ----AGNANANANANANAGAGGGGYGGLGS-QGTSGRGGLGGQ-GAGANANANANAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 1306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
34.6%; Score 514; DB 2; L.
Best Local Similarity 42.0%; Pred. No. 9.6e-24;
Matches 131; Conservative 20; Mismatches 125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genetics:
Gene: Rv0578c
:Superfamily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      266 GAGQGGYGGL--GSQGTSGIRRP 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               748 NGGAGGAGGSPGGSSGIOGNMGP 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GYGGLGSQGTSG 282
                                                                                                                                 26
                                                                                                                                                                                                      514
                                                                                                                                                                                                                                                                                                                                                                                                                                157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  271
                                                         셤
                                                                                                                              ò
                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
: Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A. Haure 393, 537-544, 1998
A. Anthors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A. Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A. Reference number: A70500; MUID: 98295987; PMID: 9634230
A. Reference number: A70696
A. Status: preliminary; nucleic acid sequence not shown; translation not shown
A. Molecule type: DNA
A. Residues: 1-853 <CCL>
A. Cross-references: GB-AL021897; CB: AL123456; NID: 93256022; PIDN: CAA17207.1; PID: e125197
C. Genetics: A. Company of A. Status and A. St
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hypothetical glycine-rich protein Rv2634c - Mycobacterium tuberculosis (strain H37RV) (Species: Mycobacterium tuberculosis from the complete genome Mycobacterium tuberculosis from the complete genome Mycobacterium proberculosis from the complete genome Mycobacterium tuberculosis from tuberculosis from tuberculosis from tuberculosis from 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ماریدی اور 183280225; GB:AL123456; NID:g3242265; PIDN:CAB02341.1; PID:e266390; A:Experimental source: strain H37Rv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  293 GSGGH----ALLWGAGGAGGAGGAGGSTGGAGGSTAGAGGAGGGGGGTGGLLFGNGGAG 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GQGA----GAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQG-----AGAAAAAAA 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            111 AAGGAGGGYGGL-----GSQGTSGRGCLGGQGAGAAAAAAAAAAAGGAGGGYGGLGSQ 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       523 AAGGAGGQAGGGGNGGNGGNGGNGGNGGNGATGGWLYGNGGAGGQGATAGAGGAGANGV 582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 GSHHHHHHGSMASGRGGLGGQ-GAGAAAAAAAAAAAGGAGQGGYGGLGSQG--TSGRGGLG 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----GR 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       218 AAAAAAAAGG-----AGQGGYGGLGSQGTS----GRGGLGGQGAGAAAAAAAA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 2; Length 853;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 778;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14 ASGRGGLGGQGA--GAAAAAAAAAGG-----AGQGGYGGLGSQGTS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 35.0%; Score 519.5; DB 2; Length Best Local Similarity 41.7%; Pred. No. 3.4e-24; Matches 134; Conservative 14; Mismatches 122; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34.8%; Score 516.5; DB 2;
42.1%; Pred. No. 4.7e-24;
tive 9; Mismatches 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           263 -AAGGAGGGYGGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Superfamily: unassigned collagens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Residues: 1-778 <COL>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    136;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

13;

211

Db 1201 GFGGDGLDASSG 1212

Search completed: December 18, 2002, 16:21:57 Job time : 12.5982 secs

This Page Blank (Uspto)

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
```

OM protein - protein search, using sw model

December 18, 2002, 16:06:04; Search time 4.46662 Seconds (without alignments) 2702.183 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-490-291-2 1485 1 MRGSHHHHHHGSMASGRGGL.....YGGLGSQGTSGIRRPAAKLN 291

Scoring table:

112892 seqs, 41476328 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

112892 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_40:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		Description	P19837 nephila cla	P03211 epstein-bar	053553 mycobacteri		petunia				Q10637 mycobacter1				Q50615 mycobacteri	P54320 mus musculu		Q50630 mycobacteri				P10495 phaseolus v	-	P18165 mus musculu		P07916 gallus gall		P15502 homo sapien				S	m	867	093484 oncorhynchu
SUMMERTES		TD.	SPD1_NEPCL	EBN1_EBV	YZ08_MYCTU	SPD2_NEPCL	GRP1_PETHY	YQ34_MYCTU	FBOH_BOMMO	GRP2_PHAVU	YD25_MYCTU	WA22_MYCTU	GRP_ARATH	Y747_MYCTU	YI18_MYCTU	ELS_MOUSE	Y278_MYCTU	YP91_MYCTU	ELS_RAT	YA68_MYCTU	YK98_MYCTU	GRP1_PHAVU	YI40_MYCTU	LORI_MOUSE	ELS_BOVIN	ELS_CHICK	LORI_HUMAN	ELS_HUMAN	SANT_PLAFV	FSH_DROME	CSP_PLACG	VNUA_PRVKA	DDX9_MOUSE	CSP_PLACC	CA21_ONCMY
		DB	7	7	Н	-	-	1	Н	-	-	-	-	П	П	-	-	-	Н	-	-	7	_	_	-	-	, i	_	-	Н	-	-	-		H
		Length	747	641	. 1901	627	384	778	5263	465	603	914	338	801	498	860	957	543	864	463	491	252	515	481	747	750	316	730	375	2038	401	1733	1380	398	1356
dБ		Match	68.4	41.1	37.5	37.3	35.4	34.8	34.2	33.4	33.4	33.3	33.0	33.0	32.9	31.8	31.4	31.2	30.5	29.9	29.9	29.5	28.2	27.6	56.6	26.0	25.5	25.2	24.4	24.1	23.0	22.4	22.3	22.1	21.9
		Score	1015	610.5	556.5	554	525	516.5	508.5	496.5	496	494	490	490	489	472	467	463	453.5	444.5	443.5	438	418.5	409.5	394.5	385.5	378	374	363	357.5	341.5	332.5	330.5	328	325
	Result	8	г	7	m	4	S	9	7	80		10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	58	30	31	32	33

1010545

34 311 20.9 1147 1 MYSB_ACACA P19706 acanthamoeb 35 305 20.5 419 1 CSP_PLACM P08676 plasmodium 36 304.5 20.5 189 1 GRP2_ONYSA P29834 Oryza sativ 37 297.5 20.0 385 1 GRP2_ONYSA P51992 xenopus lac 38 293 19.7 1355 1 CA21_RANCA CA2350 rana catesb 39 292.5 19.7 1355 1 CA21_RANCA CA2350 rana catesb 40 290 19.7 1372 1 CA21_RANCA CA2350 rana catesb 41 288.5 19.7 1372 1 CA21_RAT P30754 riftia pach CA2 287 19.3 1372 1 CA21_RAT P30754 riftia pach 44 284 19.2 1262 1 GA44_CAEEL CA21_CAAEL CA32_CA32 canis famil 45 283.5 19.1 156 1 GA44_CAEEL CA32_CA33 cacnorhabdi 45 283.5 19.1 156 1 GA42_CAEEL CA32_CA32_CA33 cacnorhabdi 45 283.5 19.1 404 1 CA2_DROME CA32_CA33_CA32_CA33_CA32_CA33_CA32_CA33_CA32_CA33_CA32_CA33_CA32_CA33_CA32_CA32	SPDI_NEPCL STANDARD; PRT; 747 AA. SPDI_NEPCL STANDARD; PRT; 747 AA. O1-RB-1991 (Rel. 17, Created) DT 01-NOV-1995 (Rel. 32, Last sequence update) DT 15-JUN-2002 (Rel. 41, Last senoncation update) DE Spidroin 1 (Dragline slik fibroin 1) (Fragment). OS Nephila clavipes (Orb spider). OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneacomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila cox Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila	RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE. RA MEDILINE-9018499; PubMed-2402494; RA Xu M., Lewis R. V.; RT "Structure of a protein superfiber: spider dragline silk." RI "Structure of a protein superfiber: spider dragline silk." RI [2] RP SEQUENCE OF 653-747 FROM N.A. REDINE-94165058; PubMed-8120021; RA MEDLINE-94165058; PubMed-8120021; RA Beckwitt R., Arcidiacono S.; RT "Sequence conservation in the C-terminal region of spider series (Spidroin) from Nephila clavipes (Tetragnathidae) RT Araneus bleenearius (Araneidae)."; RL J. Biol. Chem. 259: 6661-6663(1994). CC characteristics of strength and elasticity. Fibroin con pseudocristaline regions of antiparallel beta-sheet in with elastic amorphous segments. CC SUBGNIT: MAJOR SUBUNIT, WITH SPIDROIN 2, OF THE DRAGLII CC SUBGNIT: MAJOR SUBUNIT. WITH SPIDROIN 2, OF THE DRAGLII CC SUBGNIT: MAJOR SUBUNIT. WITH SPIDROIN 2, OF THE DRAGLII CC SUBGNIT: MAJOR SUBUNIT: WITH SPIDROIN 2, OF THE DRAGLII CC SUBCELLUAR LOCATION: Extracellular. CC DOMAIN: Highly repetitive protein characterized by reg. CC DOMAIN: Highly repetitive protein characterized by reg.	CC This SWISS-PROT entry is copyright. It is produced through a collabor between the Swiss Institute of Bioinformatics and the EMBL outsits the European Bioinformatics Institute. There are no restrictions of use by non-profit institutions as long as its content is in n modified and this statement is not as long as its content is in n modified and this statement is not send in email to license agreement (See http://www.isb-sib.ch/ann.cc entities requires a license libense in the sib.cc. in the EmBL; N33137; AAA29380.1; EMBL; N37137; AAA29380.1; EMBL; N0N TER 1

```
ð
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                 13;
                                                                                                          NPGLSGCDVLIGALLEVVSALIQILGSSSIGQVNYGSAGQA
TQIVGQSVYQAL -> ILVFLDVMSSFKLFSRLFLLLSRS
(IN REF. 1).
850E44B0D649E012 CRC64;
                                                                                                                                                                                    GGQGAGAAAAAA...........AAAAGGAGQGGYGGLGSQGTSGRGGLG 135
                                                                                                                                                                                                                       136 GQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA---- 182
                                                                                                                                                                                                                                                                          385 GCYGGLGSQGAGRGGQGAGAAAAAGGAGQRGYGGLGNQG-AGRGGLGGQGAGAAAAAA 443
                                                                                                                                                 Gaps
                                                                                                                                                           ------AAAAAAGGAGGGKG 45
                                                                                                                                                                               96
                                                                                                                                                                                                         GLGSQGTSGRGGLGGQGAGAAAAAAAAA------AGCAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                 -------AAAAAGAGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA--
                                                                                                                                                                                                                                            14; Indels 142;
                                                                                                                                     Score 1015; DB 1; Length 747;
Pred. No. 7.6e-44;
8; Mismatches 14; Indels 14:
                                                                                                                                                                                                                                                                                                                                                              <del>4</del>
                                                                                                                                                                                                                                                                                                                                                   Epstein-barr virus (strain B95-8) (Human herpesvirus Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Lymphocryptovirus. [1]
                                                                                                                                                                                                                                                                                           444 GGAGQGGYGGLGNQGAGRGGQGAAAAAGGAGQGGYGGLGSQG 485
                                                                                                                                                                                                                                                                                     -> L (IN REF. 1).
-> T (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                  21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                          11 GSMASGRGGLGGQGAGAAAA------
                                                                                                                                                                                                                                                                                                                        PRT;
          9.
111.
112.
114.
116.
116.
                                                                                                                            ¥.
                                                                                                                                      68.48;
                                                                                                                                           59.28;
                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                  EBNA-1 nuclear protein.
747 AA;
                                                                                                                                           Best Local Similarity
Matches 238; Conserv
21-JUL-1986
21-JUL-1986
                                                                                            REPEAT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                         EBN1_EBV
                                                                                                                                      Query Match
                                                                  REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
              REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                              REPEAT
REPEAT
                                                        REPEAT
REPEAT
                                                                                                                                                                              46
                                                                                                                                                                                                                                           183
                                                                                                                                                                                                                                                                                     260
                                                                                                                                                                                                   97
8
                                                                                                                                                                              ö
                                                                                                                                                                                       g
                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                     ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10;
MEDLINE-84270667; PubMed-6087149;
Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
Tuffnell P.S., Barrell B.G.;
"DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
Nature 310:207-211(1984).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RGSHHHHHHGSMASGRGGLG-GQGAGAAAAAAAAAAAAAGGAGQ-GGYGGLGSQGTSGRGGLG 59
                                                                                                                                                                                     [2]
SEQUENCE OF 1-26 FROM N.A.
MEDLINE-86259739; PubMed-3460083;
MEDLINE-86259739; PubMed-3460083;
MEDLINE-86259739; PubMed-3460083;
"Nucleotide Sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: a probable transcriptional initiation site.";
Proc. Natl. Acad. Sci. U.S.A. 83:5096-5100(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-90266473; PubMed-2161150;
Petti L., Sample C., Kieff E.;
Subunclear localization and phosphorylation of Epstein-Barr virus latent infection nuclear proteins.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -1- FUNCTION: INVOLVED IN LATENT CYCLE. EBNA-1 FUNCTIONS IN THE MAINTENANCE REPLICATION OF EBV EPISOME. TRANSACTIVATING FACTOR FOR THE ORGIN AND ENHANCER FUNCTIONS OF ORIP.
-1- SUBCELLULAR LOCATION: NUCLEAR. FREE IN THE NUCLEOPLASM, SOMEWHARSCOTATED WITH THE CHROMATIN AND HARDLY, IF AT ALL ASSOCIATED WITH THE NUCLEAR MATRIX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nuclear protein; DNA-binding; Transcription regulation; Activator;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          60 GQGAGAAAAAAAAAAAAAGGAG-GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAG---GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                116 GQGGYGGLGSQGTSGRGGLGGQ-GAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       domain of the Epstein-Barr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 641;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-96006523; PubMed-7553871;
Bochkarev A., Barwell J.A., Pfuetzner R.A., Furey W.F.
Edwards A.M., Frappier L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLY/ALA-RICH.
4D161653E16FC341 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 469-607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41.1%; Score 610.5; DB 1
48.6%; Pred. No. 3.8e-24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative 14; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Crystal structure of the DNA-binding of virus origin-binding protein EBNA 1."; Cell 83:39-46(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  352 G
56427 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; V01555; CAA24816.1; -. EMBL; M13941; AAA45889.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR004186; EBNA1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF02905; EBNA1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDB; 1VHI; 23-DEC-96
TRANSFAC; T00211; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; A03773; QQBE31
PIR; S33021; S33021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              641 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 143; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
```

us-09-490-291-2.rsp

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14;
                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.

STAIN-H37RV;

MEDLINE-98295987; PubMed-9634230;

A Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F., Bordock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gencles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Ollorer S., Osborner J., Quall M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Sequres R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

"Deciphering the blology of Mycobacterium tuberculosis from the Nature 393:537-544(1998).
175 QGAGAAAAAAAAAGGAGQGGYGGLGSQGTSG----RGGLGGQGAGAAAAAAAAGGAG 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65 AAAAAAAAAAAAAGGAGGGY----GGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAG-QGG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        120 YGGL----GSQGTSGRGCLGGQGAGAAAAAAAAGGAG-----GGYGCLGSQGTSGR 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            416 AGGAGGVGGOGGNAIAGGINGSGGAGGTGGQGGAGGMGGSGADNASGIGADGGAGGTGGN 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Mycobacterium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HYPOTHETICAL PE-PGRS FAMILY PROTEIN
                                                                      284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 1901;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --
--
--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS SUBFAMILY.
                                                                                          231 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGIR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Pred. No. 3e-21;
13; Mismatches 112; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C7B1923D5D0146CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hypothetical protein; Repeat; Signal; Complete proteome.
SIGNAL 1
                                                                                                                                                                                                                                                    (Rel. 39, Last sequence update)
(Rel. 40, Last annotation update)
PE-PGRS family protein Rv3508 precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1;
                                                                                                                                                                                                 PRT; 1901 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    37.5%; Score 556.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RV3508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1901 AA; 147627 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tuberculist; Rv3508;
InterPro: IPRO00084; PE_region.
Pfam; PF00934; PE; E;
ProDom; PD001223; PE_region; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             _
=
-
-
-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; AL022022; CAA17745.1; -.
                                                                                                                                                                                                                                 30-MAY-2000 (Rel. 39, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     44.98;
                                                                                                                                                                                                                                                                                                                              Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                         RV3508 OR MTV023.15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-1773;
                                                                                                                                                                                                                                                                                          Hypothetical
                                                                                                                                                                                                                                                      30-MAY-2000
                                                                                                                                                                                                                                                                     16-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 140;
                                                                                                                                                                                             YZ08_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local
                                g
                                                                  ŏ
                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                   Hinman M.B., Lewis R.V.;
"Isolation of a clone encoding a second dragiline silk fibroin.
Nephila clavipes dragiline silk is a two-protein fiber.";
J. Biol. Chem. 267:19320-19324(1992).
-!- FUNCTION: Spiders major ampullate silk possesses unique characteristics of strangth and elasticity. Fibroin consists of pseudocristalline regions of antiparallel beta sheet interspersed
                                               652
                                                                                         653 GAGGAIGTGGTGGAVGSVGNAGIGGTGGTGGVGGAGGAGAAAAAGSSATGGAGFACGAGG 712
                                                                                                                                                                                                                                                                                                                                                                                                                                536 AGGAGGAGGNTGVGGTNGSGGQG-GTGGAGGAGGAGGVGADNPTGIGGTGGTGGKG--GA
                        GCLGGQG--AGAAAAAAAAAGG-AGQGGYGGLGSQGT---SGRGGLGGQGAGAAAAA
                                                                     AAAGGAGQGGYGG-----LGSQGTSGRGGLGGQ-GAGAAAAAAAAAAAGGA----GQGG
                                                                                                                                                                                                                                                                    Nephila clavipes (Orb spider).
Eukaryota; Metazoa: Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15 APPROXIMATE TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ٠, :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CB9B63779B2C594B CRC64;
                                                                                                                                                                                                                     01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Spidroin 2 (Dragline silk fibroin 2) (Fragment).
                                                                                                                                                                                                   ₹
                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-92406876; PubMed-1527052;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              55.
70.
10.
111.
14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54184 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M92913; AAA29381.1; -. Silk; Repeat.
                                                                                                                272 YGGLGSQGTSGI 283
                                                                                                                                       713 EGGAG--GNSGV 722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ž
                                                                                                                                                                                                                                                                                                        NCBI_TaxID=6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SPD2_NEPCL
P46804;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
REPEAT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                              593
                                                                   224
                                           g
                                                                   ð
                                                                                         g
                                                                                                                                      8
                                                                                                                                                                                                             ò
                                                                                                                à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license also removed. (See http://www.isb-sib.ch/announce/or send an email to license@isp-sib.ch).
                                        17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
SUBCELLULAR LOCATION: Cell wall (Potential).
MISCELLANEOUS: THIS PROPEIN CONTAINS 67% GLYCINE RESIDUES.
MISCELLANEOUS: 90% OF THE MATURE PROTEIN RESIDUES ARE CAPABLE OF PORMING A BETA-PLEATED SHEFT COMPOSED OF 8 ANTI-PARALLEL STRANDS.
MISCELLANEOUS: THE GLYCINE-RICH REGION IS COMPRISED OF TWO RELATED FAMILIES OF REPEATS, F1 AND F2, EACH REPEAT CONTAINING ABOUT 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Condit C.M., Meagher R.B.; "A gene encoding a novel glycine-rich structural protein of petunia."; Nature 323:178-181(1986).
                                                                                                                                                                                                                                                                                                                                                                                                 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                        274
                                                                                               GTSGRG----GLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRG----GLGGQGAG 102
                                                                                                                                                                                       147 GPGGYGPGQQGPSGPGSAAAAAAAAASGPGQQGPGGYGP-GQQGPGGYGPGQQGPSGPGSA 205
                                                                                                                                                                                                                          206 AAAAAAASGPGQQGPGGYGP-GQQGPCGYGPGQQGLSGPGSAAAAAAAAAGPGQQGPGGYGP 264
                                                                                                                                                                                                                                                                                                                                     265 GOOGPSGPGSAAAAAAAAAGPGGYGPGQQGPGGYGPGQQGPSGAGSAAAAAAAGPGQQGL 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       383 SAAAAAAAAGPGGYGPGQQGPGGYAPGQQGPSGPGSASAAAAAAAAGPGGYGPGQQGPGG 442
                                        Gaps
                                                                          3 GSHHHHHHGSMASGRGGLGGQGAGAAAAAAAAAGGAGQ---GGYG-----GLGSQ 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnollophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
                                                                                                                                                                                                                                                                                                    145 ------AAAAAAGGAGGGYG-CLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                          194 GGYGGLGSQGTSGRG----GLGGQGAGAAAAAAAA-GGAGQGGYGGLGSQGTSGRG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                    247 -----GLGGGGAG-----GAGQGGYGG
                                        94;
 Length 627;
                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAR-1989 (Rel. 10, Created)
01-MAR-1989 (Rel. 10, Last sequence update)
01-AUG-1992 (Rel. 23, Last annotation update)
Glycine-rich cell wall structural protein 1 precursor.
                                      101;
   Score 554; DB 1;
Pred. No. 2.1e-21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; X04335; CAA27866.1; -.
PIR; A26099; A26099.
Cell wall; Structural protein; Repeat; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             384 AA
                                        Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
37.3%; Scott 44.1%; Pretive 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Petunia hybrida (Petunia).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            275 L--GSQGTSGIRRPAA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 443 YAPGOOGPSGPGSAAA 458
                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=4102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GRP1_PETHY P09789;
   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ä.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Cell wa
                   Best Loca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GRP1_PETHY
                                                                                                                                                    21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                             g
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                       쉽
                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                    ö
                                                                                                                                                                                     8
                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
```

```
ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S.L.,
                                                                                                                                                                           60 SFGRGRGAGGGFGGGAGGGAGGGLGGGGGGAG--GGGGLGGGGGGGGGFGGGAGG 117
                                                                                                                                                                                                                             231
                                                                                                                                                                                                                                                                                                                                                                          238 GGGVGGGVGGGAGGGGGGGGGGGGGGGLGGGSGHGGGFGAGGGVGGGAAGGVGGGGF 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     s:
                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W., Umayam A.A., Ermolaeva M.D., Salzborg S.L. Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                              Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Harris
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Osborne J., Osborne J., Morlean J., Moule S., Murphy L., Oliver S., Seeger K., Skelton S., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; Deciphering the biology of Mycobacterium tuberculosis from the Nature 393:537-544(1998).
                                                                                                                                                                                                            AAAGG-AGQGGYGGLGSQGTSGRGGLG-------GQGAGAAAAAAAAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                          120 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGS----QGTSGRGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome comparison of Mycobacterium tuberculosis clinical
                                                                                                              22;
GLYCINE-RICH CELL WALL STRUCTURAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
                                                                               Length 384;
                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                        232 GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLG 276
                                                                                                                                                                                                                                                                                                                                                                                                                                         GGGGGGGVGGGSGHGGGFGAGGGVGCGAGGGLGGGGGGGGGGIG 342
                                               C8541C549417D18C CRC64;
                                                                                              Pred. No. 4.2e-20;
2; Mismatches 136;
                                                                               DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical Per-PGRS family protein Rv2634c.
RV2634C OR MT2712 OR MTCY441.04c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        778 AA
                                                                             Score 525;
                                GLY-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-98295987; PubMed-9634230;
                            384 G
28777 MW;
                                                                            35.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-CDC 1551 / Oshkosh;
                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  laboratory strains.";
                                               384 AA;
                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=1773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-H37RV
                                                                                             Best_Local Simi
Matches 125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YQ34_MYCTU
P71933;
                                             SEQUENCE
                                                                               Query Match
                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 6
YQ34_MYCTU
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                       298
SHH
                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                          ð
```

ហ

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                  16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGLGGQGAGAAAAA--AAAAGGAGQGGYGG--LGSQGTSGRGGLGGQGAGAAA--AAAA 109
                                                                                                                                                                                                                                                                                                                                                                                                                     213 Q--GAGAAA-----AAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAG 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            110 AAAGGAG------QGGYGGLGS-QGTSGRGGLGGQGA----GAAAAAAAAAAAAGGAGQG 156
                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                              ASGRGGLGGQGA--GAAAAAAAAAAGG------AGQGGYGGLGSQGTS-----GR 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pancrustacea; Hexapoda;
Lepidoptera; Glossata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GYGGLGSQGTSGRGGLGGQ----GAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGCLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Esnault C.,
Li Z.-G.;
gene.";
                                                                                                                                                                                                                                                                                      DB 1; Length 778;
                                                                                                                                                                                                                                                                                   Similarity 42.1%; Pred. No. 1.6e-19; Sonservative 9; Mismatches 122; Indels
                                                                                                                                                                                                               51 51 V -> L (IN REF. 2).
63 63 Q -> H (IN REF. 2).
274 274 A -> T (IN REF. 2).
778 AA; 63131 MW; DAB20FE58B4999E7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDIANE-2333352; Pubmed-10871375; Zhou C.-Z., Confalonieri F., Medina N., Zivanovic Y., Yang T., Jacquet M., Janin J., Duquet M., Perasso R., "Fine organization of Bombyx mori fibroin heavy chain Nucleic Acids Res. 28:2413-2419(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FBOH_BOMMO STANDARD; PRT; 5263 AA. P05790; Q26379; Q17220; Q1001-NOV-1988 (Rel. 09, Created) 16-CCT-2001 (Rel. 40, Last sequence update) 16-CCT-2001 (Rel. 40, Last annotation update) Fibroin heavy chain precursor (Fib-H) (H-fibroin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota, Metazoa, Arthropoda, Mandibulata,
Insecta, Pterygota, Neoptera, Endopterygota,
Ditrysia, Bombycoidea, Bombycidae, Bombyx.
                                                                                                              GAGQGGYGGL--GSQGTSGIRRP 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           748 NGGAGGAGGSPGGSSGIQGNMGP 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-168 FROM N.A. MEDLINE-80045039; Pubmed-498286;
                                                                                                    EMBL; Z80225; CAB02341.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bombyx mori (Silk moth).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=7091;
                                                                                                                                                                                                                                                                                                                  Matches 136;
                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                      Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FBOH_BOMMO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FIBH.
                                                                                                                                                                                                                                                                                                                                                                                                     26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 7
    õ
                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its conțent is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-9939630; Pubbed=10366732;
Tanaka K., Kajiyama N., Ishikura K., Waga S., Kikuchi A., Ohtomo K.,
Takagi T., Mizuno S.;
Takagi T., Mizuno C., Takagi T., Takagi 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PARTIAL SEQUENCE FROM N.A. MEDILINE-94356842; PUBMOG-7916056; MEDILINE-94356842; Pubmod-7916056; Mita K., Ichimura S., James T.C.; "Highly repetitive structure and its organization of the silk fibroin
Tsujimoto Y., Suzuki Y.;
"The DNA sequence of Bombyx mori fibroin gene including the 5'
fibaning, mRNA coding, entire intervening and fibroin protein coding
                                                                                                                                                                                                                                                                                                                                                                     STRAIN-Kinshu X Showa;
MEDLINE-89094868; PubMed-3210244;
Mita K., Ichimura S., Zama M., James T.C.;
"Specific codon usage pattern and its implications on the secondary structure of silk fibroin mRNA.";
J. Mol. Biol. 203:917-925(1988).
                                                                                                                                                                                     MEDLINE-79211211; PubMed-455439;
Tsujimoto Y., Suzuki Y.;
"Structural analysis of the fibroin gene at the 5' end and its surrounding regions.";
Cell 16:425-436(1979).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INTERCHAIN (WITH LIGHT CHAIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 5179-5263 FROM N.A., AND DISULFIDE BONDS STRAIN-J-139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FIBROIN HEAVY CHAIN.
HIGHLY REPETITIVE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -> V (IN REF. 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL, AF226688; AAF76683.1;

EMBL; V00094; CAA23433.1;

EMBL; S74439; AAB3161.1;

EMBL; X13869; CAA32076.1;

EMBL; M3789; CAA32076.1;

EMBL; M3789; CAA32076.1;

EMBL; M37818, AAA27839.1;

PIR; S01844; S01844.

SIGNAL

SIGNAL

22 S263 FIBRC
                                                                                                                                       [3]
PARTIAL SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           J. Mol. Evol. 38:583-592(1994).
                                                                                                                                                                                                                                                                                                                                             PARTIAL SEQUENCE FROM N.A.
                                                                                      regions.";
Cell 18:591-600(1979)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
DISULFID
DISULFID
CONFLICT
```

PROTEIN 1.8.

ä

```
YD25_MYCTU
Q10637;
               DOMAIN
DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YD25_MYCTU
     FFFS
                                                                                                                                                   g
                                                                                                                                                                                                     a
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                            음
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@18b-sib.ch).
                                                      12;
                                                                                                    4939 GYGAGAGYGAGYGYGAGAGYGAGAGSGAGSGAGAGSGAGAGSGAGAGSGAGAGS 4998
                                                                                                                                                                                                                                                                      5059 GSGAGAGSGAGGYGAGGSGAASGAGAGSGAGAGAGAGAGSGAGAGSGAGAGYGAGA 5118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
-1- SUBCELLIARA LOCATION: Cell wall (Potential).
-1- DOMAIN: THE REPEATED DOMAINS OF THE PROTEIN FORM A BETA-PLATED SHEET CONFIGURATION.
                                                                                                                                   230 GQG-----GYG-GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQG---GYG-GL 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JUL-1989 (Rel. 11, Last sequence update)
01-OCT-1994 (Rel. 30, Last annotation update)
Phassolus vulgaris (Kidney bean) (French bean).
Spermatophyta; Magnoliophyta; eudloctyledons; core eudlocts; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
                                                      Gaps
                                                                              11 GSMASGRGGLG-GQGAGAAAAAAAAAGGAGQG-----GYG-GLGSQGTSGRGGLGGQGA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY: THE NATIONAL SIGNAL SEQUENCE OF ABOUT 30 AA SHOWS ABOUT 60% HOMOLOGY TO THAT OF THE GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN GRP 1.0.
                                                                                                                                                                                                                                           172 LGGQGAGAAAAAAAAAAGGAGG-GYG-GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGA
                                                      37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
GLYCINE-RICH CELL WALL STRUCTURAL
                            Length 5263;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Keller B., Sauer N., Lamb C.J.; "Glycine-rich cell wall proteins in bean: gene structure and
                         Query Match 34.2%; Score 508.5; DB 1; Length 5 Best Local Similarity 43.6%; Pred. No. 1.2e-18; Matches 134; Conservative 16; Mismatches 120; Indels
 8EE11D3A0A47440E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               association of the protein with the vascular system."; EMBO J. 7:3625-3633(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   wall; Structural protein; Repeat; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  465 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-cv. Tendergreen;
MEDLINE-89091109; Pubmed-3208742;
 391586 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUL-1989 (Rel. 11, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; X13596; CAA31932.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
 5263 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S01820; S01820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                               5119 GSGAASG 5125
                                                                                                                                                                                                                                                                                                                                                      276 GSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=3885;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 GRP2_PHAVU
P10496;
SEQUENCE
                                                                                                                                                                                         119
                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
တ္တ
                                                                               á
                                                                                                                                   ð
                                                                                                                                                             9
                                                                                                                                                                                       à
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                              셤
```

```
SEQUENCE FROM N.A.
STRAIN=CDC 1551 / Oshkosh;
STRAIN=CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bishai W.;
"Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                 251 GGAGGGQGGGAGGGYGAGGEHGGGAGGGQGGGAGGCYGAGGEHGGGAGGGQCGGAGGGYG 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               311 AGGEHGGGGGGGGGGGGGGGAAAVGEHGGGYGGGQGGGDGGGYGTGGEHGGGYGGGQGGG 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        135 GEHGIGYGGGGSGAGGGGGYNAGGAQGGGYGTGGGAGGGGGGG----GDHGCGYGGGQG 190
                                                                                                                                                                                                                                                                                                                                                                             63 AGAAAAAA-----AAAAGGAGQGG--YGGLGSQGTSGRGGLGGQGAGAAAAAAA 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             s:
                                                                                                                                                                                        35; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-H37RV;

MEDLINE-98295987; PubMed-9634230;

A Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D Cole S.T., Brosch R., Parkhill J., Cannier T., Churcher C., Harris D Gordon S.V., Elglmeier K., Barn D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S. A Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., A Oliver S., Osborno J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

T "Deciphering the biology of Mycobacterium tuberculosis from the Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.NCBI_TaxID-1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Harris
                                                                                                                                                                                                                                                   3 GSHHHHHHGSMASGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    163 SQCTSGRGGLGGQGAGAAAAAAAA----AGGAGQGGYGGLGSQGTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       110 -AAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA-----AAAGGAGQGGYGGLG
                                                                                                                         Length 465;
                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
GLY-RICH.

8 X 22 AA TANDEM REPEATS.
B5C4A9B983B43607 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv1325c precursor.
RV1325C OR MT1367 OR MTCX130.10C.
                                                                                                                         DB 1;
                                                                                                                                                    Pred. No. 1.2e-18;
8; Mismatches 139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       603 AA
                                                                                                                         Score 496.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                       36683 MW;
                                                                                                                      33.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mycobacterium tuberculosis.
                                                                                                                      Query Match 33.4'
Best Local Similarity 41.3'
Matches 128; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                    465 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      431 GG-GAGGGGG 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           273 GGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
```

^

```
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          202
 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Biolinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                      16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---TSGRGGLGGQGAGAAAAAAAAAGGAGQ--------GGYGGLGSQ 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           285 GAGGAGGFGVSGSAGGNGGTGGDGGIFTGNGGTGGTGGTGTGNQLVGGEGGAGGAGGNAG 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---GQGGYGGLGSQ------GTSGRGGLGGQGAGAAAAAAAAAAGGAG------ 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGGIGGSGGSGGAAKLIGDGGAGGTGGDSVRGAAGSGGTGGTGGLIGDGGAGGAGGTGIE 463
                                                                                                                                                                                                                                                                                                                                                                                GAGGAGGDGPGNGAAGGAGGAGGLLFGSGGAGGPGGVGNTGTGGLGGDG-GAAGLFGAGG 225
                                                                                                                                                                                                                                                                                                                                                                                                     AGGAG----QGGYGGLGSQ----------GTSGRGGLGGQGAGAAAA 107
                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                           16 GRGCLGCQCAGAAAAAAAAAAGGA--GQGGYGGLGSQCTSGRGCLGCQCAGAAAAAAAA 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QGGYGGLGSQGTS----GRGGLGGQGA----GAAAAAAAAAAAGG-AGQGGYGGLGSQG--
                                                                                                                                                                                                                                                                                                                                                                                                                                              108 AAAAAGGAG----QGGYGGLGSQ----GTSGRGGLGGQGAGAAAAAAAAAAAAAGGA----
                                                                                                                                                                                                     POTENTIAL.
HYPOTHETICAL PE-PGRS FAMILY PROTEIN
                                                                                                                                                                                                                                                                                                                Score 496; DB 1; Length 603;
Pred. No. 1.4e-18;
9; Mismatches 118; Indels 100;
SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                 InterPro: IPR000084; PE_region.
Pfam; PF00934; PE; 1.
ProDom; PD001223; PE; 1.
Hypothetical protein; Repeat; Signal; Complete proteome.
SIGNAL 1 30
                                                                                                                                                                                                                                               G -> V (IN REF. 2).
G -> D (IN REF. 2).
G -> D (IN REF. 2).
G -> F (IN REF. 2).
M; 4F9BCBB2B07AE964 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    914 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WAG22 antigen precursor.
WAG22 OR RV1759C OR MT1807 OR MTCY28.25C.
Mycobacterium tuberculosis.
                                                                                                                                                                                                                            RV1325C.
GLY-RICH.
                                                                                                                 EMBL; Z73902; CAA98089.1; -.
EMBL; AE007010; AAK45630.1; ALT_INIT.
TIGR; MT1367; -.
                                                                                                                                                                                                                                                                                           49575 MW;
                                                                                                                                                                                                                                                                                                                Query Match
33.4%;
Best Local Similarity 37.6%;
Matches 137; Conservative 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                Tuberculist; Rv1325c; -.
                                                                                                                                                                                                                                      114
132
135
337
508
603 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        279 GTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GDGG 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WA22_MYCTU
006794;
                                                                                                                                                                                                                                                                      CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
WA22_MYCTU
                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              524
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                              Harris D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     403
                                                                                                                                                                                                                                                                                                                                                                ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           164 GNGGAGGAGTATANNGIGGAGGAGGSAMLFGAGGAGGAGGAATSLVGGIGGTGGTGGNA 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66 AAAAAAAAAGGAG-----QGGYGGLGSQG----TSGRGGLGGQG-AGAAAAAAAAA-- 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      284 GAGGMGGAGGFGDHGTLGTGGAGGDGGGGLFGAGGDGGAGGSGLTTGGAAGNGGNAGTL 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----QGTSGRGGLGGQG--AGAAAAAAAAAAAGGAGGGGYGGLGSQG--TSGRGGLGGQG 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gapa
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                MEDLINE-9629997; PubMed-9634230;

Gole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris Gordon S.V., Elgimeler K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Deviln K., Feltwell T., Gentles S., Hamlin N., Holroyd (Horsby T., Jagels K., Kroph A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quall M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Mitchead S., Barrell B.G.;

"Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  224 GMLAGAAGAGGAGGFSFSTAGGAGGAGGAGGLFTTGGVGGACGQGHTGGAGGAGGAGGLF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16 GRGGLGGQGAGAAAAAAAAAAGGA------GQGGYGGLGSQGTS---GRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----QGGYGGLGSQGTS----GRGGLGGQGAGAAAAAAAAAAGGAG-----QGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SLGAAGGAGGTGGAGGTVFGGGKGGAGGAGGNAGMLFGSGGGGGTGGFGFAAGGQGGVGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGAGQGGYGGLGSQGTSGRGLGGQGAGAA-AAAAAAAAGGAG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 85.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY, PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 914;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            33.3%; Score 494; DB 1; Length 914; 39.4%; Pred. No. 2.2e-18; Live 10; Mismatches 126; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WAG22 ANTIGEN.
F6953C3DBE8E6AC8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00934; PE; 1.
Antigen; Repeat; Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AE007040; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TubercuList; Rv1759c; -.
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE 914 AA; 74354 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; Z95890; CAB09322.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 134; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TIGR; MT1807;
```

```
RESULT 12
Y747_MYCTU
                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ibs-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .
6
                                                                                                                                                                                                                                                                                              gene from Arabidopsis thaliana.";
Plant Mol. Biol. 17:949-952(1991).
-!- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
-!- SUBCELLULAR LOCATION: Cell Wall (Potential).
                                                                                                                          01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 23, Last annotation update)
01-A0G-1992 (Rel. 23, Last annotation update)
Arabidopsis thallana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
118 IGGGAGGGSGGCLGGGIGGGAGGGGGCGLGGGIGGGGAGGGAGGGLGGGGIG 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                  STRAIN=cv. Columbia; MEDLINE=92003708; PubMed=1912511; Quigley F., Villiot M.L., Mache R.; "Nucleotide sequence and expression of a novel glycine-rich protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11 GSMASGRGGLGGQGAGAAAAAAAAAGGAG---QGGYGG--LGSQGTSGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 178 -- GGAGGGSGGLGGGIGGGAGGGAGGGGAGGGLGGGHGGGFGGGAGGGLGGGAGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         185 AAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 33.0%; Score 490; DB 1; Length 338; Best Local Similarity 42.4%; Pred. No. 2e-18; Matches 118; Conservative 10; Mismatches 140; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLY-RICH.
046A6E8C1A4E89EB CRC64;
                            245 RGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          294 GGVGGFFGGSGGFFGGGAGGGAGGGFGGGGGGAG 331
                                         EMBL; X58338; CAA41249.1; -.
PIR; S17732; KNMU.
Cell wall; Structural protein; Repeat; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23891 MW;
                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21 3
338 AA;
                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21
                                                                                                        GRP_ARATH P27483;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                               GRP_ARATH
      g
                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRAIN=CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŝ
                                                                                                                                                                                                                                   Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";
Submitted (ARR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Harris
                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-9892987; PubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harriss Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harriss Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd Burnsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Seeger K., Krogh A., McHean J., Moule S., Murphy L., Suston J.E., Taylor K., Whitehead S., Barrell B.G.;
Deciphering the biology of Mycobacterium tuberculosis from the Complete genome sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
HYPOTHETICAL PE-PGRS FAMILY PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Repeat; Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SSING (IN REF. 2).
EA54C9BF45A00F41 CRC64;
                                                      30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv0747 precursor.
Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     G -> S (IN REF. 2).
R -> G (IN REF. 2).
K -> R (IN REF. 2).
S -> G (IN REF. 2).
S -> G (IN REF. 2).
A -> P (IN REF. 2).
T -> A (IN REF. 2).
      801 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MISSING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Tubercutist, Rv0747; -.
InterPro: IPR000084; PE_region.
Pfam: PF00934; PE; 1.
Hypothetical protein; Repeat; S
SIGNAL 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AL021958; CAA17514.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65407 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AE006968; AAK45011.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             577
580
801 AA;
                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TIGR; MT0772.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBFAMILY
Y747_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bishai W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                              053810;
```

Length 801;

DB 1;

33.0%; Score 490;

Query Match

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                    14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P. STRAIN-H37RY;
(X. MEDLINE-98295987; PubMed-9634230;
(X. MEDLINE-98295987; PubMed-9634230;
(X. MEDLINE-98295987; PubMed-9634230;
(X. MEDLINE-98295987; PubMed-9634230;
(X. MEDLINE-98295987; Parkhill J., Garnier T., Churcher C., Harris D.,
(A. Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
(A. Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
(A. Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
(A. Badcock K., Basham D., Chillingworth T., Connor R.,
(A. Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
(A. Badcock K., Basham D., Chillingworth T., Connor R.,
(A. Badcock K., Basham D., Chillingworth T., Connor R.,
(A. Badcock K., Basham D., Chillingworth A., McLean J., Moule S., Murphy L.,
(A. Badcock K., Skelton S., Squares R.,
(A. Skelton S., Seeger K., Skelton S., Squares R.,
(A. Batham J., Chillingworth T., Chillingworth J.,
(A. Batham J., Chillingworth M., Robert J.,
(A. Batham 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ernolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                ----AGAAAAAAAAAAAGG------AGQGGYGGLGSQGTSGRGGLGGQGAGA 103
                                                                                                                                                                                                                                                                                                                               104 AAAAAAAAAGG----AGQGGYGGLGSQ-GTSGRGGLGGQGAGAAAAAAAAAAAGG----AG 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                      191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LLFGSGGAGGSGGFSNSGNGGAGGAGGAGLVGSGGAGGAGASATGAATGGDGGAG--G 719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        s:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                  11 GSMASGRGGLGGQGAGAAAAAAAAAGGA----GQGGYGGLGSOGTSGRGGLGGOG---
                                                                                                                                                                                                                                                                                                                                                                   155 QGGYGGLGSQGTS-----GRGGLGGQ--------GAGAAAAAAAAAAAAGGA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---GQGGYGGLGSQGTSGRGGLGGQGAGA-----AAAAAAAAAAGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
Pred. No. 3.3e-18;
4: Mismatches 120; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     720 KSGAFGLGGDGGAGGATGLSGAFHIGGKGGVGGSAVLIGNGGNGG 764
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGG----LGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
19-pothetical PE-PGRS family protein Rv1818c.
RV1818C OR MT1866 OR MTCY1A11.25C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ġ.
                              14; Mismatches
      38.68;
                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
      Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YI18_MYCTU
Q50615;
                                    133;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bishai W.;
                                                                                                                                                   425
                                                                                                                                                                                                             63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99 QCAGAAAAAAAAAAGGAGQGGYGGL-GSQGTSGRGGLG--GQCAGAAAAAAAAAAGGAGQ 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           292 VGGDGVAFLGTAPGGPGGAGGAGGLFGVGGAGGAGGIGLVGNGGAGGSGGSALLWGDGGA 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGYGGLGSQGTSGRGGLGGQ-----GAGAAAAAAAAAAAAGGAGGGYGGLGSQGTS----G 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            173 GNVASGTAGFGGAGGAGGLIYGAGGAGGAGGRAGGGVGGIGGAGGAGGNGGLLFGAGGAG 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11 GSMASGRGGLGGQGA-----GAAAAAAAAAGGAGQGGYGGLGSQGTS----GRGGLG 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Meiazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID-10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GCAGGVGST-TGGAGGAGGNAGLLVGAGGAGGAGALGGGATGVGGAGGNG--GTAGLLFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RGGLGGQGAGAAAAAA--AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 GQGAGAAAAAAAAAAGG-----AGQGGYGGLGSQGTS------GRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 498;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        32.9%; Score 489; DB 1; Length 49 ilarity 40.1%; Pred. No. 2.8e-18; Conservative 14; Mismatches 123; Indels
                                                                                                                                                                                                  proteome.
                                                                                                                                                                                                                                                                                                                                                                                               MISSING (IN REF. 2).
G -> GAGG (IN REF. 2).
4F6F78F2482586BA CRC64;
                                                                                                                                                                                                  Complete
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
Elastin precursor (Tropoelastin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       860 AA
                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      256 AAAAAAAAGGAGQGGYGG-LGSQGTSGI 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                469 GNAGSGIPNGSAGTGGAGGLLGKNGMNGL 497
                                                                                              EMBL; Z78020; CAB01461.1; -. EMBL; AE007045; AAK46139.1; ALT_INIT.
                                                                                                                                                                                                   Transmembrane;
                                                                                                                                   Tuberculist; Rv1818c; --
Tuberculist; Rv1818c; --
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE, 1.
ProDom; PD001223; PE_region; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-OCT-1996 (Rel. 34, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                              40755 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                Hypothetical protein;
TRANSMEM 1 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                              498 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                          FIGR; MT1866;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Simi
Matches 132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ELS_MOUSE
P54320:
                                                                                                                                                                                                                                                                                                TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                              FRANSMEM
                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                        FRANSMEM
                                                                                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                             TRANSMEM
                                                                                                                                                                                                                                                                                                                                             TRANSMEM
                                                                                                                                                                                                                                                                                                                                                          FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                      PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
TIGR; MT0291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bishai W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      688
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15;
                                                                                                                                               Chromosome 7.";
Genomics 23:125-131(1984).

- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND
- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND
- SUNCTION: WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.
- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER
INTO AN EXTENSIBLE 3D NETWORK.
- SUBCELLUIAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.
- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLG--SQGTSGRGGLGGQGAGA- 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------SGRGGGY-GGLGSQGT----SGRGGL 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GIGGIAGAGTPAAAAAAAAAAKAAKYGAAGGLVPGGPGVRLPGAGIPGVGGIPGVGGIPG 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16 GRGGLGGQGAGAAAAAAAAGGA------------------GQGGYGG 46
                                                                                                             Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           646 FGAGAGVPGFGAGAVPGSLAASKAAKYGAAGGLGGPGGLGGPGGLGGAGVPGRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                466 AGLGGASPAAAAAAAKYGAGGAGALGGLVPGAVPGALPGAVPAVPGAGGVPGAGTPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGQGAGAAAAAAA------AAAG-------GAGQGGYG-GLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAAAAAAKAAAKAGLGPGVGGVPGCVGVGGIPGGVGVGGVPGGVGPGGVTGIGAGPGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches 100; Indels 128;
                                    STRAIN-BALB/C; TISSUE-Lung;
MEDINE-95130069; PubMed-7823060;
Wydner K.S., Sechler J.L., Boyd C.D., Passmore H.C.;
"Use of an intron polymorphism to localize the tropoelastin mouse chromosome 5 in a region of linkage conservation with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; DB 1; Length 860; 2.6e-17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPWU059,9, TROPOELASTIN.
PRINTS; PR01500; TROPOELASTIN.
Structural protein; Repeat; Signal; Connective tissue.
77 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OCOBESAAE1EDD7F1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            253 AG----AAAAAAAAAAGGAGQGGYGGLGSQGTSGI 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31.8%; Score 472; 38.5%; Pred. No. 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MGD; MGI:95317; Eln.
InterPro; IPR003979; tropoelastin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                71955 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; U08210; AAA80155.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 152; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (Rel. 39, 0
(Rel. 39, 1
(Rel. 40, 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                860 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P56877;
30-MAY-2000 (
30-MAY-2000 (
16-OCT-2001 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Y278_MYCTU
P56877;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    526
SO THE TANK 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseelsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Godson M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                       STRAIN=H37kv;

MEDLINE=98295987; PubMed-9634230;

A Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., A Cole S.T., Brosch R., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Badcock K., Brogia R., Morphy L., A Davles R., Dayles K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

Boctphering the biology of Mycobacterium tuberculosis from the Complete genome sequence.;

I Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GCLGSQGTSGRGCLGGQGAGAAAAAAAAGGAGQGGYGGLGSQ----GTSGRGGL---G 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGNG--GLFANGGAGGAGGFNAAGGNGGNGGLFGTGGTGGAGTNFGAGGNGGNGGLFGAG 745
                                                              Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGWLFGNGGNGGHGATNTAATATGGAGGAGGILFGTGGNGGTGGIAT-GAGGIGGAG-GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11 GSMASGRGGLGGOGAGAAAAAAAAGGA----GOGGYGGLGSOGTSGRGGLGGOGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          66 AAAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGA---GAAAAAAAAAAGG--AGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGVSLLIGSGGTGGNGGNSIGVAGIGGAGGRGGDAGLLFGAAGTGGHGAAGGVPAGVGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HYPOTHETICAL PE-PGRS FAMILY PROTEIN RV0278C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12; Mismatches 129; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Tuberculist; Rv0278c; -.
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE; 1.
Hypothetical protein; Repeat; Signal; Complete protecome.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        M -> I (IN REF. 2).
MISSING (IN REF. 2).
R -> G (IN REF. 2).
71EBABD417FBA47C CRC64;
Hypothetical PE-PGRS family protein Rv0278c precursor. RV0278c OR MT0291 OR MTV035.06C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31.4%; Score 467; DB 1; 39.7%; Pred. No. 4.8e-17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AL021930; CAA17353.1; -.
EMBL; AE006936; AAK44511.1; ALT_INIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81905 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-CDC 1551 / Oshkosh; Fleischmann R.D., Alland D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40
163
807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30
957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40
158
807
957 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                        NCBI_TaxID=1773
```

Search completed: December 18, 2002, 16:17:54 Job time : 9.46662 secs

OKEN ANDRO ORDERINA

ĕ

```
084199 mycobacters
096418 mycobacters
096418 mtheraeap
084120 mycobacters
085359 mycobacters
053559 mycobacters
05355 mycobacters
05357 mycobacters
063135 mycobacters
084121 mycobacters
084121 mycobacters
083135 mycobacters
083135 mycobacters
083118 mycobacters
083118 mycobacters
083118 mycobacters
083118 mycobacters
083118 mycobacters
083118 mycobacters
083318 mycobacters
08300 mycobacters
08300 mycobacters
08300 mycobacters
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           006818 mycobacter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [1]
SEQUENCE FROM N.A.
MEDIJNE-94165058; Dubmed-8120021;
BECKWILT R., Arcidiacono S.;
"Sequence conservation in the C-terminal region of spider silk profeins (Spidforin) from Nephila clavipes (Tetragnathidae) and Araneus bicentenarius (Aranejae).;
J. Biol. Chem. 269:6661-6663(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-98318805;
Beckwitt R., Arcidiacono S., Stote R.;
Beckwitt R., Arcidiacono S., Stote R.;
"Evolution of repetitive proteins: spider silks from Nephila clavipes terragnathidae) and Araneus bicentenarius (Araneidae).";
Insect Biochem. Mol. Biol. 28:121-130(1998).
EMBL; U37520; AAC04504.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     68.7%; Score 1020; DB 5; Length 617;
55.3%; Pred. No. 4.8e-53;
Live 6; Mismatches 12; Indels 178; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Dragline silk protein spidroin 1 (Fragment).
Nephila clavipes (Orb spider).
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae; Araneomorphee; Entelegynae; Araneoldea; Tetragnathidae; Nephila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    49410 MW; 1EF0CE9269A832E2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ALIGNMENTS
                                                     08VIX9
09BHT7
08VIZ0
053559
09BHU6
053557
00BHU6
053557
016986
08VIZ3
08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     053394
09GUB5
053890
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             006818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09V5V8
Q9BIU1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 68.7
Best Local Similarity 55.3
Matches 242; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                       617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           617 6
617 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=6915;
567.5
563.5
563
563
560.5
557
556
556
                                                                                                                                                                                                              525.5
525.5
5224.5
5204.5
512.5
512.5
512.5
500.5
500.5
503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         046172 046172;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SETTER TO A REPORT OF THE PROPERTY OF THE PROP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         046171 nephila cla
09biv1 argiope aur
09biv1 argiope tri
09biu3 dolomedes tr
09biu4 dolomedes tr
09biu4 dolomedes tr
09biu4 nephila cla
017434 nephila cla
01698 araneus dia
09bit4 latrodectus
09bit5 artheraea pr
09bit6 antheraea pr
09bit6 nephila mad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             antheraea p
antheraea y
nephila mad
pinctada fu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               046172 nephila cla
                                                                                                                                                                                                              (without alignments)
3509.535 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                               December 18, 2002, 16:13:04; Search time 17.0848 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                                                                                                                                                                                1485
1 MRCSHHHHHGSMASGRGGL.....YGGLGSQGTSGIRRPAAKLN
                      GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  671580 segs, 206047115 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QBWSW4
Q46171
Q9BIU7
Q9BIU7
Q9BIU7
Q9BIU6
Q9BIU6
Q17434
Q1698B
Q9BIT9
Q9BIT9
Q9BIT9
Q9BIT9
                                                                                                                                                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SPTREMBL_21:*

1: sp_archea:*
2: sp_bacteria:*
3: sp_fung1:*
4: sp_human:*
5: sp_lnvertebrate:*
6: sp_mammal:*
7: sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_organelle:*
sp_phage:*
sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0% Maximum Match 100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB
                                                                                                                                                                                                                                                                                            US-09-490-291-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              544
447
648
6691
854
854
10
988
410
253
373
2639
2655
2633
738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6688.7
6688.7
555.1
555.1
555.1
555.1
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.5
668.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1020
1017
1018
8118.5
8118.5
689.5
689.5
689.5
640.5
640.5
617.5
617.5
617.5
617.5
617.5
                                                                                                                                                                                                                                                                                                                     Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                      Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       rotal number
                                                                                                                                   protein
                                                                                                                                                                                                                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database
                                                                                                                                                                                     ë
:
```

δy	50 QGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAA 109	qq	284 GL
q	: 107 QG-AGRGGSGGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGGGAGAAAAAA 160	6y	183
οy	110 AAAG	qq	343 LG
q	161 GGAGDGGYGGLGGQGGGGGGGGGGGGGGGGGGGGGGAGAAAAAGGAGGGGGG	Qy	231 06
δλ	145AAAAAAGGGGGGGGGGGGGT	đ	398 QG
Q	220 AGQGAGAAAAAAGGAGGGYGGLGSQGAGRGGQGAGAAAAAVGAGQGGYGGQGAGQGGY 279	Qy	259 AA
δ	167SGRGGLGGQGAGAAA	qu	458 AA
g	280 GGLGSQGAĞRGĞLĞĞĞĞAĞAAAAAAAGGAGQGGLGGQGAGQGAGAAAAAAAGGAGQGGYGG 339	<u> </u>	c E
ογ	183AAAAAAGGAGGGGGGGGGGGGGGGGGAAAAAAAA	046171	KESULT 3 046171 TD 046171
q	340 LGNQCAGRGCQCAAAAAGGACQGGYCGLGSQG-AGRGCLGGQGAGAAAAAAGGAGQGGY 398	3 8 5	046171; 01-mm.
ογ	226AGGACQGGYGCLGSQGTSGRGCLGGQGACAAAAAA 261	555 	01-JUN-
a à	399 GGLGGQGAGQGGYGGLGSQG-SGRGGLGGQGAGAAAAAGGAGQGGLGGQGAGQGAGAAA 457 262 aaaacaacaacaa 226	OS	Spidroi Nephila
전	AAAGGVRQGGYGGLGSQG	888	Eukaryo Araneom NCBI Tay
RESU	LT 2	RN	SEQUENCI
O8WS OID	QBWSW4 PRELIMINARY; PRT; 644 AA.	RR	Xu M.,]
25.5	Olyman-2002 (TrEMBLrel, 20, Created)	RL RN	Proc. Na [2]
3 5 5	OI-WAR-2002 (TEMBLEEL: 20, Last sequence update) OI-JUN-2002 (TEMBLEEL: 21, Last annotation update) Dragline silk nrotsin (France)	RP RX	SEQUENCI MEDLINE
00 C		RT	Arcidiac "Purific
888	bunaryota, metazoa, mrumopoua; cherrerata; Aradamida; Aradeae; Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila. NCBI TaxID-6915;	RI	Appl. M
R P	[1] SEQUENCE FROM N.A.	RP RX	SEQUENCE MEDLINE
8 5	Basel R.; Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.	RA	"Evolut
g g	<pre>EMBL; AY061814; AAL32375.1; InterPro; IPR000104; Antifreeze_1.</pre>	RT	(Tetragr Insect
Y E	PRINTS; PRO0308; ANTIFREEZEI. NON_TER 644 644	RP	[4] SEQUENCE
og og	SEQUENCE 644 AA; 52104 MW; 9024AF2F65C5F75A CRC64;	RA	Beckwitt Submitte
2 8 2	544;	DR DR	EMBL; UZ
è	CACCARA TION TO THE TO THORE TO GE	SO	SEQUENCE
g 43	50 GSQGAGRGGLGGQGAGAAAAAAAGGGGGGGGGGGGGGGGGGG	Õ m s	Query Match Best Local
δi	50 QGT		_
e è		qq	 64 GSC
G 6	110 AGGAGOGGLGGQCACAAAAAAGGAGCCCCCCCCCCCCCCCCCCCCC	QY	52 TSC
δλ	110 AAAGGAGGGGGGGGGGGGGGGGGGGGGAGAAAAAAAAA	q ₀	1
qq		λο i	
ογ	160 GLGSQGTSGRGGLGGQGAGAAA	qq —	178 AG

```
13;
t R., Arcidiacono S., Stote R.;
ion of repetitive proteins: spider silks from Nephila clavipes
nathidae) and Araneus bicentenarius (Araneidae).";
Biochem. Mol. Biol. 28:121-130(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  146 | 146 | 147 | 147 | 147 | 147 | 148 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 
                                                                                                                                                                                                                      NQGAGRGGQGAAAAAAGGGGGGGGGGGGGGA---AAAAAAGGGG 397
                                                                                                                                                                                                                                                                                                                                                             GYGGLGSQGT-----SGRGGLGGQGAGAAA---------- 258
                                                                                                                                                                                                                                                                                                                                                                                                                                  67.9%; Score 1008; DB 5; Length 544;
69.9%; Pred. No. 2.2e-52;
Live 8; Mismatches 17; Indels 74; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ASGRGGLGGQGAGAAAA-------AAAAAAGGAGQGGYGGLGSQG 51
                                                                                                                                                           ------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1998 (TrEMBLrel. 06, Created)
1998 (TrEMBLrel. 06, Last sequence update)
2001 (TrEMBLrel. 19, Last annotation update)
n 1 (Fragment).
clavipes (Orb spider).
ta; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae; Orphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
xID-6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      E OF 449-544 FROM N.A. :=98148687; PubMed=9487707; cono S., Mello C., Kaplan D., Cheley S., Bayley H.; cation and characterization of recombinant spider silk ed in Escherichia coli.'; idrobiol. Biotechnol. 49:31-38(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ure of a protein superfiber: spider dragline silk."; atl. Acad. Sci. U.S.A. 87:7120-7124(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 t R.;
ed (JAN-1995) to the EMBL/GenBank/DDBJ databases.
20329; AAC38957.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   544 AA; 44107 MW; CC611B3551945615 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                544 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               E FROM N.A.
=90384959; PubMed=2402494;
Lewis R.V.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAAAGGAGQGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity 69.930; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         E FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -98318805
```

```
ద
                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13;
           250
                                                75 GGAGOGGYGGLGSQGT----SGRGGLGGQGAGAAAAAAAAAAAGGAGGGGYGGLGSQGTS 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96 GGAGGAGRGGLGAGGAGQGYGSGLGGQGGAGQGGAAAAAAAAAGGQGGQGGYGGLGSQG-- 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       130 GRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAG 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              190 GAGQGGYG-----GLGSQGTSGR-----216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               210 GAGEGGLGAGGAGGGYGSGLGGQGGAGGGGAAAAAAAAAGGQGGHGGYGGLGSQGAGQGGA 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA 74
                                                                                                                                                                                                                                                                                                                                                                                                                                       SQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLG----SQGTSGRGGLG----G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Arthropoda; Chellcerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Araneidae; Argiope.
NCBI_TaxID=156844;
                                                                                                                                                                                                                                                                                         Spider Silk
                                                                                                                                                                                                                                                                                                                                                                                        DB 5; Length 447;
                                                                                                                                                                                                                                                                                                                                                                                      S5.1%; Score 818.5; DB 5; Length Similarity 60.2%; Pred. No. 2.5e-41; 7; Conservative 14; Mismatches 55; Indels
                                                                                                                                                                                                                                                                              Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
"Extrame Diversity, Conservation, and Convergence of SF
Fibroin Sequences.";
Science 291.2603-2605(2001).
EMBL, AR30262; AR30591.1;
InterPro; IPR000817; Prion.
PRINTS; PR00341; PRION.
                                                                                                                                                                                                                                                                                                                                                           447 447
447 AA; 35630 MW; 78D160B7E6DF24C7 CRC64;
                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Æ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            648
                                                                          350 QGAGAAAAAAAGGVRQGGYGGLGSQG 378
                                                                                                                                                            01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequen
01-OCT-2001 (TrEMBLrel. 18, Last annota
Major ampullate spidroin 1 (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         271 -----GYG-GLGSQGTSGIRRPAA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LGAGGAGQGYGXGLGGQGGAGQRGAAS 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                             PRT;
                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. MEDLINE=21179804; PubMed=11283372;
                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                               Argiope aurantia.
                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Sin
Matches 197;
                                                                                                                                                                                                                                                                                                                                                           NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9BIU7;
                                                                                                                                             Q9BIV1
                                    201
                                                                                                                                                                                                                                                                                                                                                                                                                            15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 5
Q9BIU7
ID Q9BII
AC Q9BII
                                                                                                                                             셤
                                   οy
                                                      g
                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
```

```
260 GGYGQGGVAAAAAAAGGAGGGGGGGGGGGGGGGGGGGGGAGG 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        174 GQGAGAAAAAAAAAAAAAGG-AGQGGYGGLGSQGT--------SGRG-- 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   74 AGGAGOGGYGGLGS-QG-TSGRGGLGGQCACAAAAAAAAAAGG-AGQGGYGGLGSQGT-- 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   129 ------GAAAAAAAG 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     152 GAGQGGYGGLGSQGT----SGRGGLG 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------GLGGQG----AGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15 SGRGGLGGQGAGAAAAAAAAAAGG-AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. MEDLINE=21179804; Pubmed=11283372; Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.; Catesy J., Hayashi C., Motriuk D., Woods J., Lewis R.; Extreme Diversity, Conservation, and Convergence of Spider Silk Fibroin Sequences."; Science 291:2603-2605(2001).
                                                                                                                                                                                                                                                                                                              Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
"Extreme Diversity, Conservation, and Convergence of Spider Silk Fibroin Sequences.";
Science 291:2603-2605(2001).
EMBL; AF350266; AAK30595.1; -.
                                                                                                                                                         Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Araneidae; Argiope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Lycosoidea; Pisauridae; Dolomedes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 648;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE 648 AA; 52887 MW; ACDB67067B2302CB CRC64;
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 55.1%; Score 817.5; DB 5 Best Local Similarity 51.8%; Pred. No. 3.8e-41; Matches 200; Conservative 14; Mismatches 45
                                                                                Major ampullate spidroin 1 (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JUN-2001 (TrEMBLrel. 17, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-21179804; PubMed-11283372;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUN-2001 (TrEMBLrel. 17, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Fibroin 2 (Fragment). Dolomedes tenebrosus.
                                                                                                                                 Argiope trifasciata.
                                                                                                                                                                                                          NCBI_TaxID=156845;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-156846;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09BIU3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   496
```

14;

Page

```
648
                                                                                                   202
                                                                                                                                                                                                                         528
                                                                                                                                                                                                                                                                  231 GGAGSGRGGYGDQGGLGGYGQGAGAGAAAAGGGDGYGQGYYGDQG--GRGGYGQGSGA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  57
                                                  -----GQGAGAAAAAAAAG-GAGQGGY----GGLGSQGT-----SGRG
                                                                                                                                                                                             160 -----GLGSQGTSGRGGLG--GQGAGAAAAAAAAA-----GGAGQGGY---GGLGSQ
                                                                                                                                                                                                                                                                                                                       236 -GLGSQGT-----SGRGGLG------GQGAGAAAAAAAAAAG-GAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SGRGGLG------GQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (1)
SEQUENCE FROM N.A.
MEDLINE-21179804; Pubmed-11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
"Extreme Diversity, Conservation, and Convergence of Spider Silk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota, Métazoa, Arthropoda, Chelicerata, Arachnida, Araneae,
Araneomorphae, Entelegynae, Araneoidea, Theridiidae, Latrodectus,
NCBL_TaxID=156851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 45.8%; Score 680; DB 5; Length 360; Best Local Similarity 57.3%; Pred. No. 2.8e-33; Matches 176; Conservative 12; Mismatches 44; Indels 75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OEB6E46EC1122740 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
Major ampullate spidroin 1 (Fragment).
                                  GAAAAAAAAA----GGAGQGGY---GGLGSQGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Fibroin Sequences.";
Science 291:2603-2605(2001)
EMBL, AF350273; AAK30602.1;
InterPro; IPR000104; Antifreeze_1.
InterPro; IPR000817; Prion.
PRINTS; PR00308; ANTIFREEZEI.
PRINTS; PR01228; EGGSHELL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         360 AA; 30253 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Latrodectus geometricus.
                                                                                                                                                                                                                                                                                                                                                           270 GGYGGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                               649 GGYGNQGGRGSYG 661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                       001E60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9BIU0
                                  64
                                                                                                                                                                                                                                                  203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12
                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
Q9BIU0
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                          g
                                                                                                            g
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                          ò
                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                          18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23;
                                                                                                                                                                                 GGLGGQGAGAAAAAAAAAAGGAGGYGGLGSQGT-------SGRGGLG 135
                                                                                                                                                                                                                                                                                                                              136 GQ------GAGAAAAAAAAAAG-GAGQGGYGGLG------SQ 164
                                                                                                                                                                                                                                                                                                                                                                                  GTSGRGGLG--GQGAGAAAAAAAAAGGAGQGGYGGLG------SQGTSGRGGL 210
                                                                                                                                                                                                                                                                                                                                                                                                               464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 524
                                                                                                                                                                                                                                                                                                                                                                                                                                       GGQGAGA-AAAAAAAAAGGAGQGGYG---GLGSQGTSGRGGLG--GQGAGAAAAAAAA 264
                                                                                                                                          Gaps
                                                                                                                                                                                                                     53 -----SGRGGLGGQ-----GAGAAAAAAAAAAG-GAGQGGYGGLGSQGTSGR 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GSMASGR-----GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG-GQGA 63
                                                                                                                                                                11 GSMASGRGGLG------GQGAGAAAAAAAAAG-GAGQGGYGGLGSQGT-----
                                                                                                                                                                                                                                                                                                                                                                                                  silk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Araneae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 46.4%; Score 689.5; DB 5; Length 854; Best Local Similarity 45.5%; Pred. No. 1.5e-33; Matches 197; Conservative 12; Mismatches 61; Indels 163;
                                                                                                                                        Indels 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Arane.
Araneomorphae; Entelegynae; Lycosoidea; Pisauridae; Dolomedes
                                                                                                               Length 691;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
SEDGING-21179804; PubMed-11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
"Extreme Diversity, Conservation, and Convergence of Spider !
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       854
71881 MW; 5C52F1B0AE31C6C2 CRC64;
                                                                                   CB46EF9F1991F33E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Fibroin 1 (Fragment).
Dolomedes tenebrosus.
                                                                                                            ;; Score 768.5; DB 5;
;; Pred. No. 3e-38;
17; Mismatches 51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Elbroin Sequences.";
Science 29:12603-2605(2001).
EMBL: AF350269; AAR30598.1; ..
InterPro; IPR000217; Tubulin.
PROSITE; PS00227; TUBULIN; UNKNOWN_1.
EMBL; AF350270; AAK30599.1; -.
InterPro; IPR00104; Antifreeze_1.
InterPro; IPR002951; Eggaball.
PRIWTS; PR00308; AWTIFREEZEI.
PRINTS; PR01228; EGGSHELL.
                                                                                  691 AA; 56176 MW;
                                                                                                           Guery Match 51.8%;
Best Local Similarity 51.7%;
Matches 196; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          G-GAGQGGYGGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   525 GAGAGQGGYGGQGGLGGYG 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   854 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=156846;
                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9BIU4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9BIU4
                                                                                                                                                                                                                                                                         94
                                                                                                                                                                                                                                                                                                                                                                                  165
                                                                                                                                                                                                                                                                                                                                                                                                            407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11
    SO TE SO
                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                 ò
                                                                                                                                                                                                                     ç
                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                       음
                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
```

17;

```
12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-98200471; PubMed-9541398;
MEDLINE-98200471; PubMed-9541398;
Colgin M.A., Lewis R.V.;
"Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like 'spacer regions'.";
Protein Sci. 7:667-672(1998).
EMBL; AF027735; AAC14589.1; -.
EMBL; AF027735; PR00341; Prion.
                               227
                                                                                                                                                         283
                                                                                                                                                                                                                                                                               271
108 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTS 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91 -----SGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGR 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      178 GAAAAAAAAAGGAGGGYGG------LGSQGTSGRGGLG---GQGAGAAAAA 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10 HGSMASGRGGLGGQGAGA-AAAAAAAAAGGGGGYG-------GLGSQGTS 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAA
                                                                                                                                                                                                                                            GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGG-AGQGGYG---GLGSQGTSGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 GRGGLG-GQGAGAAAAAAAAAAG------GAGQGGYGGLGSQGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     223 AAAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
NCBI_TaxID=6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 675; DB 5; Length 988;
Pred. No. 1.2e-32;
6; Mismatches 95; Indels 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE 988 AA; 79082 MW; 461E03DF53F7085D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           017434 PRELIMINARY; PRT; 988 AA.
017434;
01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
Minor ampullate silk protein MiSpl (Fragment).
Nephila clavipes (Orb spider).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 45.5%;
Best Local Similarity 45.4%;
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        598 YGGQGGYG 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                     284 RRPAAKL 290
                                                                                                                                                                                                                                                                                                                                                                                                                               AAAASAL 265
                                                                                                                                                                                                                                                                                                                                                                                                                                  259
                                                                                                                      168
                                                                                                                                                                                                                                               228
                                                                                                                                                                                                                                                                                                         208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
017434
AC 01744
DT 01-7
DT 01-7
DT 01-7
DD 01-7
DD MINO
OS Neph
OC Aran
OX NCBI
RN [1]
RP SEQU
RX MEDL
RA COLG
RT "Spil

                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                  ò
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 :|::|||||||| :| ||||| || || || || || SGSSAAAAAAA--SGPGGYGPGGYGPGSGGYGPGSGGAGGPGGPGASAAAAAA 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             112 AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGG 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    172 LGGQGAGAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQ 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       183 YGPGGPGSSAAAAAA --GSGPGGYGP-ENQGPSGPGGYGPGGSGSSAAAAAAAA--SGP 237
                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE=96178678; Pubmed=8600519;
Guerette P.A., Ginzinger D.G., Weber B.H., Gosline J.M.;
"Silk properties determined by gland-specific expression of a spider fibroin gene family.";
Science 272:112-115(1996).
EMBL; U47856; AAC47011.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                69
                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                               14 ASGRGGLG--GQG------AGAAAAAAAAAAGGACQGGYGGLGSQGTSGRGGLGGQG 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGYGGLGSQGTSGRGGLGGQ-----GAGAAAAAAAAAAGGAGGGYGGLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Araneus diadematus (Spider).
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Araneidae; Araneus.
NCBI_TaxID=45920;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S118
                                                                                                                                                                                                                                                                                                                                                                                                                  43;
                                                                                                                                                                                                                                                                                                                                                                                43.1%; Score 640.5; DB 5; Length 410; 56.9%; Pred. No. 6.4e-31; tive 24; Mismatches 61; Indels 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDINE-21179804; PubMed-11283372; Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.; Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.; Extreme Diversity, Conservation, and Convergence of Spider Fibroin Sequences."; Science 291.12603-2605(2001). Embl., AF350279; AR430608.1; -. PRINTS; PRO1608; BACINVASINC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nephila senegalensis.
Eukaryota, Metazoa, Arthropoda, Chelicerata, Arachnida,
Araneomorphae, Entelegynae, Araneoidea, Tetragnathidae,
NCBL_TaxID-156852;
                                                                                                                                                                                                                                                                                                                                                  SEQUENCE 410 AA; 34904 MW; E87A77BD59FFF35C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
Major ampullate spidroin 1 (Fragment).
                                                                           Last sequence update)
Last annotation update)
                                 Ź
                                                              Created)
                                                      Ol-OCT-2000 (TrEMBLrel. 01, La 01-OCT-2000 (TrEMBLrel. 15, La Fibroin-4 (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 56.9%
Matches 169; Conservative
                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                   NON TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9BIT4
Q9BIT4;
RESULT 10
Q16988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                         53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122
                                                                                                                                       RESULT
                      g
                                           ò
                                                              g
                                                                                   ŏ
                                                                                                      g
                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                              10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17;
                                                                                                                        79 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG 138
                                                                                                                                                                 139 AGAAAAAAAAAAGGAGGGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGG 198
                                                                                                                                                                            106 -----AAAAAAAGGAGGGYG-GLGSQGTSGRG-----GLGGQGAGAAAAAAAAAGGA 153
                                                                Gaps
                                                                                                     1 GLGGGGAGRGAGAAAAAGGAGGGGY------GGLGGQGAG----AAAAAAGGAG 45
                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                19 GLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAG
                                                                                                                                                                                                      199 IGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14 ASGRGGLG-----GQGAGAAAAAAAAAAGGAGQGGYG-GLGSQGTSGRG-----GLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Latrodectus geometricus.
Bukaryota, Metazoa, Arthropoda, Chellcerata, Arachnida, Araneae;
Araneomorphae, Entelegynae, Araneoidea, Theridildae, Latrodectus.
NCBL_TaxID-156851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Silk
                                                              16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Spider
                                          Length 253;
                                                              12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
"Extreme Diversity, Conservation, and Convergence of Sp
Fibroin Sequences.";
                      9D6B1D34A69B9869 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          373
29724 MW; ED8E0E2E16CA59E8 CRC64;
                                                                                                                                                                                                                                                                                                                                                   01-JUN-2001 (TrEMBLrel. 17, Last sequence update) 01-JUN-2002 (TrEMBLrel. 21, Last annotation update) Major ampullate spidroin 2 (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 5;
                                       Score 634; DB 5;
Pred. No. 1.1e-30;
5; Mismatches 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 632.5; DB 5
Pred. No. 1.8e-30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17; Mismatches
                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fibroin Sequences.";
Science 291:2603-2605(2001).
EMBL; AF350274; AAK30603.1; -
INTERPRO; IPRO00104; ANTIFFEEZEI.
PRINTS; PR00308; ANTIFFEEZEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-21179804; Pubmed-11283372;
 1 1
253 253
253 AA; 21311 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 52.1%;
Matches 170; Conservative 1
                                        Query Match
Best Local Similarity 61.7%;
Matches 150; Conservative
                                                                                                                                                                                                                                                                                                                                          01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   373 AA;
                                                                                                                                                                                                                                                259 AAA 261
                                                                                                                                                                                                                                                                   165 AAS 167
NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                     Q9BIT9
                                                                                                                                                                                                                                                                                                 RESULT 12
                                                                                                                                                                                                                                                                                                             Q9BIT9
                                                                                                                                           쉽
                                                                                                                                                                                                                          g
                                                                                                    g
                                                                                                                         à
                                                                                                                                                                 ð
                                                                                                                                                                                   음
                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                              ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            а
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                å
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1322 GSSAAAAAAAAAAAARRAGHDRAAGSAAAAAAAAAAAAASGAGGSGGGYGWGDGGYGSD 1381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             137 --QGAGAAAAAAAAAAAGAGQGGYGGLGSQGT------SGRGGLGG----- 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-20496953; Pubmed-11040284;
MEDLINE-20496953; Pubmed-11040284;
Sezutau H., Yukuhiron K.;
"Dynamic rearrangement within the antheraea pernyi silk fibroin gene is associated with four types of repetitive units.";
J. MOL. Evol. 51:329-338 (2000).
BENBL; AF083334; AAC32606.1; -.
SEQUENCE 2639 AA; 216057 MW; 2EE3310DEEB0999A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             213 QGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAG-GAGQGG 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pteryyota; Neoptera; Endopterygota; Lepidoptera; Glóssata; Ditrysla;
Bombycoldaa; Saturnildae; Saturniinae; Saturniini; Antheraea.
NCBI_TaxID=7119;
154 GQGGYG----GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGY--GGLGSQGTSGR
                                            179 GRGGYGPGAAGPGNYGPSGPGGSG-----AAASAAAASGPGGQQGYGPGGSGAAAAAAS
                                                                                          208 GGLG------GQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              175 -----QGAGAAAAAAAAAAGGAGQGGYGGLGS-------QGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 2639;
                                                                                                                                                                                                                                                                                                                                                                                  08, Created)
08, Last sequence update)
19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             41.6%; Score 617.5; DB 5;
39.8%; Pred. No. 6.2e-29;
iive 10; Mismatches 90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Antheraea pernyi (Chinese oak silk moth)
                                                                                                                                                                                                            261 AAAAGGAGQGGYG----GLGSQGTSG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1562 DGGYGS-GSSAAAAAA 1577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YGGLGSQGTSGIRRPAA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1998 (TrEMBLrel.
01-NOV-1998 (TrEMBLrel.
01-DEC-2001 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Simi
Matches 174;
```

```
9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       108 AAAA----AGGAGQGGYG-------GLGSQGTSGRGGLGGQGAGAAAA 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       146 AAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA--GGAGQGGYGGLGSQG 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               204 T-----SGRGGLGG-----SGRGGLGG 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      48 GSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA 107
                                                                                                                                                                                                                                                                                                                                                   40.5%; Score 602; DB 5; Length 2655;
40.9%; Pred. No. 5e-28;
tive 9; Mismatches 106; Indels 121; Gaps
                                                                                                                              Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pteryyota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysla; Bombycoidea; Saturniidae; Saturniinae; Saturniini; Antheraea.
                                                                                                                                                                                                                                                                                                                                                                                                              TRANSPOSON=MARINER-LIKE ELEMENT;
Lee J.-S., Hwang J.-S., Kang S.-W., Suh D.-S., Jin B.-R., Kim Y.-S.
Lee K.-S., Goo T.-W., Yun E.-Y., Kim K.Y.;
"Complete nucleotide sequence of fibroin gene cloned from Antheraea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nephila madagascariensis.
Bukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
NCBL_TaxID-115969;
                                                                                                                                                                                                                                                                             yamamai.";
Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AF325500; AAK83145.1; -.
SEQUENCE 2655 AA; 218672 MW; 89BB810EF9CC4CA6 CRC64;
                                                      01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-2001 (TrEMBLrel. 17, Last sequence update) 01-JUN-2002 (TrEMBLrel. 21, Last annotation update) Major ampullate spidroin 1 (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1483 SAAAAAAAAAAASGAGGAGGYGGYGSDSAAAAAAAA 1521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Z
                            2655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUN-2001 (TIEMBLIEL: 17, Created)
01-JUN-2001 (TIEMBLIEL: 17, Last seq
01-JUN-2002 (TIEMBLIEL: 21, Last anno
                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A. MEDLINE-21179804; PubMed=11283372;
                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 40.9
Matches 163; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                   Antheraea yamamai.
                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                  Fibroin.
                           0964F4
0964F4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9BIT6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          239
RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         о9віт6
                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
79 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGAGGGGYGGLGSQGTSGRGGLGGQG 138
                                                                                                                                                                                                                                                                                                                                                                    139 AGAAAAAAAAAAAGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG 198
                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                        84
                                                                                                                                                                                                                                                                 19 GLGGQGAGAAAAAAAAAAGGAGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                       57 QGGYGGLGSQ------AGQG
                                                                                                                                                                                                                                                                                                                                                                                                                            199 LGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA
                S1.1.k
                                                                                                                                                                                                                       65;
                                                                                                                                                                                     Score 597.5; DB 5; Length 233;
Pred. No. 1.4e-28;
3; Mismatches 39; Indels 65
   Datesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Extreme Diversity, Conservation, and Convergence of Spider
Fibroin Sequences.":
                                                                                                                                                            233 AA; 19885 MW; 255ED4F5ADC94F2D CRC64;
                                                                                                                                                                                                                    13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        completed: December 18, 2002, 16:20:34
ne : 21.0848 secs
                            Fibroin Sequences.";
Science 291:2603-2605(2001).
EMBL, AF350277; AAK30606.1; -.
InterPro; IPR000104; Antifreeze_1.
PRINTS; PR0308; ANTIFREEZEI.
PRINTS; PR01608; BACINVASINC.
PRINTS; PR01504; TUBBYPROTEIN.
Hayashi C., Motriuk D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     259 AAAAAAGGAGQGGYGGLGSQG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 176 LSSTISNAVSQIGASNPGLSG 196
                                                                                                                                                                                       Query Match
Best Local Similarity 55.2%;
Matches 144; Conservative 1:
                                                                                                                                NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Search cor
   g
                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                    à
```

This Page Blank (uspto)

Nephila clavipes a Spider dragine va Polymer of an anal Spider dragine va Polymer of an anal Spider recombinant N. clavipes spidro N. clavipes spidro N. clavipes spidro Spider recombinant pmissi misp spider spider slike protein related to Spider dragine va Polymer of an anal Silvii protein com Spider dragine va Polymer of an anal Silvii protein com Silvii ke protein silvike protein silvike protein silviike silviike protein silviike protein silviike silvii

Spider dragline va Polymer of an anal

N. clavipes spidro Spider silk protei Spider natural sil Nephila clavipes s N. clavipes spider Repetitive protein Sirii protein seq Silk-like protein Amino scid sequenc Amino scid sequenc

SLP4 multimeric pr SLP4 synthetic pro

OM protein

Run on:

sednence:

Searched:

Database

Result

```
silk protein; pQE((SP1)4/(SP2)1)4; purification; fibre; spinning.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Spider recombinant silk protein pQE((SP1)4/(SP2)1)4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
                                                           AAY59070

AAR14308

AAR30105

AAR40105

AAR901178

AAR90045

AAR90045

AAM50045

AAM50046

AAM50046

AAM50046

AAM50046

AAM50046

AAR900909

AAR90940

AAR90307

AAR90307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAW26348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Butler MM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB82609 standard; Protein; 681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11-NOV-2000; 2000WO-US30086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-JAN-2000; 2000US-0490291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            02-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wello CM, Arcidiacono S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (MELL/) MELLO C M.
(ARCI/) ARCIDIACONO S.
(BUTL/) BUTLER M M.
(USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Spider; orb-weaver;
structural protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2001-483136/52.
N-PSDB; AAH26302.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO20015333-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26-JUL-2001.
                                                                                                                                                                                                                                                                                                                                      1331.5
1233.5
1233.5
1233.5
1233.5
1233.5
1233.5
1233.5
1176.5
1176.5
                                              2135
2135
2135
2126
2076
2076
1972.5
1972.5
1971.5
1711.5
1711.5
1711.5
1711.5
1711.5
1711.5
1711.5
1711.5
1711.5
1711.5
1711.5
                                                                                                                                                                                                                                                                        1425
1405.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB82609;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB82609
ID AAB8
Spider recombinant
Spider recombinant
N. clavipes spidro
N. clavipes spidro
Dragline protein 1
Dragline protein 1
Dragline protein 1
Dragline protein 1
                                                                                                                           (without alignments)
1995.696 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
4. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
5. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
7. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
9. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
11. (SIDS2/gcgdata/geneseqy-embl/AA1990.DAT:*
12. (SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
13. (SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
14. (SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
15. (SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
16. (SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
17. (SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
18. (SIDS2/gcgdata/geneseqy-embl/AA1991.DAT:*
18. (SIDS2/gcgdata/geneseqy-embl/AA1992.DAT:*
                                                                                                        December 18, 2002, 16:05:10 ; Search time 45.4697 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /SIDS2/gcgdata/geneseg/genesegp-embl/AA2000.DAT:*/SIDS2/gcgdata/geneseg/genesegp-embl/AA2001.DAT:*/SIDS2/gcgdata/geneseg/genesegp-embl/AA2002.DAT:*
                                                                                                                                                                    US-09-490-291-4
3493
1 MRGSHHHHHHGSMASGRGGL.....GGYGPGQQTSGIRRPAAKLN
              GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                  Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                      908470 seqs, 133250620 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB82610
AAM50037
AAM50039
AAM11794
AAU11797
AAU11793
AAM50042
                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A_Geneseq_101002:*
                                                                                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      333333355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     681
691
1255
1880
989
1617
1626
818
809
630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0
97.1
69.3
69.3
67.9
65.7
65.7
65.7
64.5
                                                                                                                                                                         Title:
Perfect score:
                                                                                                                                                                                                                                       Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3390
2419
2419
2372
2394
2294
2254
2250
2139
```

SQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA

541 541

> g ð

601

AAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSG

```
The present sequence is that of orb-weaver spider (Nephila clavipes)

recombinant silk protein pQE((SP1)4/(SP2)14. The invention

provides methods for purifying and spinning spider silks and other

structural proteins.

cells or other biological samples (such as non-recombinantly

cells or other biological samples (such as non-recombinantly

cells or other biological samples (such as non-recombinantly

proteins by hydrolysing many of the macromolecules while leaving

the structural proteins intect. In the case of silk proteins, the

proteins by hydrolysing many of the macromolecules while leaving

the structural proteins intect. In the case of silk proteins, the

crowding lysate is further purified by ion-exchange or affinity

choming. In the present case, the pQE((SP1)4/(SP2)1)4 gene

was cloned into vector pQE-9 for recombinant expression in

Escherichia coil, and recombinant silk protein was obtained in 97%

purity using propionic acid and anion-exchange chromatography on

QAE-Sephadex A50. Products obtained using the methods of the

including films, fibres, woven articles, sutures, ballistic

protection, parachutes and parachute cords. The new method has

the following advantages over prior art: it involves fewer steps,

requires less time and smaller volumes of reagants, results in

better recovery of protein at higher purity (70-99%), is easy to

scale up, and the fibres are spun in an environmentally benign

combined to the contraction and cost.
Recovering structural polypeptides in a biological sample, useful for purifying and spinning spider silks and other structural proteins, comprises treating the sample containing the polypeptides with an acid
                                                                                                                                                    Claim 2; Page 34-37; 49pp; English,
```

681 AA; Seguence

ö 300 61 OGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGY 120 GGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSG 180 240 300 540 AAAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSG 420 480 Gaps 9 9 RGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGG RGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGG **AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG** ö 100.0%; Score 3493; DB 22; Length 681; 100.0%; Pred. No. 1.1e-247; Live 0; Mismatches 0; Indels 0; Conservative Local Similarity 681; Query Match Matches 121 121 181 181 241 301 301 361 361 421 481 421 g ô a å g ò g à g à g à g 셤 à õ

540

481

셤

```
The present sequence is that of orb-weaver spider (Nephila clavipes)
recombinant silk protein pET((SP1)4/(SP2)1)4. The invention
provides methods for purifying and spinning spider silks and other
structural proteins. Organic acids are used to lyse recombinant
cells or other biological samples (such as non-recombinantly
derived cells), and enrich the purity and yields of structural
proteins by hydrolysing many of the macromolecules while leaving
the structural proteins intect. In the case of silk proteins, the
resulting lysate is further purified by ion-exchange or affinity
chomicoraphy and processed into an aqueous-based mixture for
fibre spinning. In the present case, the pET((SP1)4/(SP2)1)4 gene
was cloned into vector pET34 for recombinant expression in
Scherichla coli, and recombinant silk protein was obtained in
75-75% purity using propionic acid and anion-exchange chromatography
on QAE-Sephadex A50. When denaturant (3 M quanidine-HCI) was added
to the propionic acid, the recombinant protein was obtained in 80%
purity. Products obtained using the methods of the invention can
be used in the construction of many materials included.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Recovering structural polypeptides in a biological sample, useful for purifying and spinning spider silks and other structural proteins, comprises treating the sample containing the polypeptides with an acid
be used in the construction of many materials including films, fibres, woven articles, sutures, ballistic protection, parachutes and parachute cords. The new method has the following advantages
                                                                                                                                                                                                                                                                                                                          Spider; orb-weaver; silk protein; pET((SP1)4/(SP2)1)4;
structural protein; purification; fibre; spinning.
                                                                                                                                                                                                                                                                                      Spider recombinant silk protein pET((SP1)4/(SP2)1)4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Butler MM;
                                                                                                                                                                                  ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 2; Page 38-40; 49pp; English.
                                                                                                                                                                            AAB82610 standard; Protein; 691
                                                   PGGYGPGQQTSGIRRPAAKLN 681
                                                                    PGGYGPGQQTSGIRRPAAKLN 681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20-JAN-2000; 2000US-0490291.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-2000; 2000WO-US30086.
                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BUTLER M M.
US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (MELL/) MELLO C M. (ARCI/) ARCIDIACONO S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mello CM, Arcidiacono
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-483136/52.
N-PSDB; AAH26303.
                                                                                                                                                                                                                                                                                                                                                                               Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                   WO200153333-A1.
                                                                                                                                                                                                                                                    02-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                     26-JUL-2001.
                                                                                                                                                                                                                  AAB82610;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BUTL/) 1 (USSA) 1
               601
                                                   661
                                                                                   661
                                                                                                                                            RESULT 2
                                                                                                                                                               AAB82610
           g
                                                 à
                                                                                                                                                                                                                  g
```

DE10113781-A1. 13-DEC-2001

Synthetic.

```
ö
                                                                                                                                                              132
                                                                                                                                                                              250
                                                                                                                                                                                                                                   252
                                                                                                                                                                                                                                                    310
                                                                                                                                                                                                                                                             370
                                                                                                                                                                                                                                                                                                       372
                                                                                                                                                                                                                                                                                                                         430
                                                                                                                                                                                                                                                                                                                                           432
                                                                                                                                                                                                                                                                                                                                                           490
                                                                                                                                                                                                                                                                                                                                                                             492
                                                                                                                                                                                                                                                                                                                                                                                               550
                                                                                                                                                                                                                                                                                                                                                                                                               552
                                                                                                                                                                                                                                                                                                                                                                                                                                 610
                                                                                                                                                                                                                                                                                                                                                                                                                                                  612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     672
                                                                                                                                            AAAAGGAGQGGYGGLGSQGTSGRGGLGGQQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSG 130
                                                                                         Gaps
                                                                                                                           72
                                                                                                         GSMASGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAA 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOISOI
over prior art: it involves fewer steps, requires less time and smaller volumes of teagents, results in better recovery of protein at higher purity (70.99%), is easy to scale up, and the fibres are spun in an environmentally benign solution reducing hazardous waste
                                                                                                                   SQGTSGPGGYGPGQQTSGRGCLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                      AAAAAAAAAAAGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGAGGGGYGGLG
                                                                                                                                                                                                                                                   SQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA
                                                                                                                                                                                                                                                                                                                         613 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQT
                                                                                                                                                                                                                                                                                      GGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQG
                                                                                         ;
0
                                                                        Length 691;
                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N. clavipes spidroin synthetic homologue SO1SO1 protein.
                                                                     97.1%; Score 3390; DB 22;
100.0%; Pred. No. 3.7e-240;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAM50037 standard; Protein; 1255 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                    Query Match 97.19
Best Local Similarity 100.
Matches 664; Conservative
                                    accumulation and cost
                                                      691 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |||||
|SGIR 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SGIR 674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAM50037;
                                                                                                                                                            73
                                                                                                                                                                                               133
                                                                                                                                                                                                                                  193
                                                                                                                                                                                                                                                                                                       313
                                                                                                                                                                                                                                                                                                                                                                                                              493
                                                                                                                                                                                                                                                                                                                                                                                                                                                  553
                                                                                                         11
                                                                                                                          13
                                                                                                                                            7.1
                                                                                                                                                                              131
                                                                                                                                                                                                                                                                     253
                                                                                                                                                                                                                                                                                       311
                                                                                                                                                                                                                                                                                                                        371
                                                                                                                                                                                                                                                                                                                                         373
                                                                                                                                                                                                                                                                                                                                                          431
                                                                                                                                                                                                                                                                                                                                                                            433
                                                                                                                                                                                                                                                                                                                                                                                                                                 551
                                                                                                                                                                                                                  191
                                                                                                                                                                                                                                                                                                                                                                                              491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      673
                                                                                                                                                                                                                                                   251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    88888888
                                                                                                         à
                                                                                                                         g
                                                                                                                                            ò
                                                                                                                                                            셤
                                                                                                                                                                              ò
                                                                                                                                                                                              οp
                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                ò
```

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein. Comprising modules, each comprising a group of sequentially arranged oligonoucleotides, each oligonoucleotide encoding a repeating unit of a spidoroucleotide. Act of synthetic protein has at least 84% homology with the Nephila clavipes spidoron protein has at comprising and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and artificial organs; (ii) as adhesion surfaces for culturing cells; and cilis as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Compared they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high sait concentration, they are easily extracted and purified. The modular construction of the invention facilitates incorporation of additional peptide-encoding sequences, e.g. to simplify purification or modulate solubility. This sequence represents the synthetic N. clavipes spidroin-1 homologue SOISOI described in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 33;
                                                                                                                                                                                                                                                                                                                  New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QGGYGGLGSQGAGRGGLGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAG 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA------AGGAGQGGYGGLGSQ 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTSGRGGLGGQGAGAAA-------AAAAAAGGAGQGGYGGLGSQGT 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----SGPGGYGP-GQQTSGRGGLGGQGAAAA-- 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    53 -----SGRGGLGGQGAGAAAA-------------AAAAAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1255;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   40,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 23;
                                                                                                                                                                                          (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11 GSMASGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           69.3%; Score 2419; DB 23; 52.2%; Pred. No. 4.4e-169; ive 19; Mismatches 40;
                                                                                                                                                                                                                              Grosse F, Guehrs K;
                                                                                                                                                                                                                                                                                                                                                                                      Claim 22; Page 29-34; 88pp; German.
                                                                                                                                       09-JUN-2000; 2000DE-1028212, 24-OCT-2000; 2000DE-1053478
                                                                                                      21-MAR-2001; 2001DE-1013781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 52.2
Matches 581; Conservative
                                                                                                                                                                                                                              Scheller J, Conrad U,
                                                                                                                                                                                                                                                             WPI; 2002-123561/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1255 AA;
                                                                                                                                                                                                                                                                                N-PSDB; ABL61038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             nvention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                     proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    167
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
```

310

g

```
ò
                                                                                                                                                                                                                                                               유
                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                            음
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                         Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOISOISO1
                                                                                                                                                                                                                                             565
                                                                                                                                                                                                                                        599
                                                                         345
                           292
                                428 AAGGAGGGYGGLGSQGAGRGGQGGAGAAAAAGGAGGGGGYGGLGSQGAGRGGQGAGAAAA 487
                                                             547
                                                                                  GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT--SGRGGLGGQGAGAAAAAAAAA 403
                                                                                                     522
                                                                                                                                                                        544
                                                                                                                                                                                                                      369 AGOGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAA 427
                                                                                                                      ----SGRGGLGGQGAGAAA 434
                                                                                                                                                                                                    GQGGLGGQGAGQAGAAAAAGGAGQCGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYG 897
                                                        AAGGAGQGGYGGLGSQGGYGGLGSQGAGAAAAAGGAGGGGGGGG
                                                                        -----AAAAAAGGAGQGGYGGLGSQGT------SGPGGYGP-GQQTSGR
                                                                                                                                                                  AAAAAAAAAAGGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA----
                                                                                                                                                                                          -----AAAAAAGGAGGGYGGLGSOGT-------
                                                                                                                                                                                                                                                             N. clavipes spidroin synthetic homologue SO1SO1SO1 protein.
                                                                                                                                                                                                                                                                                          1076 GAGAAAAAGGAGQGGYGGLGSQG-AGRGGQGAG 1108
                                                                                                                                                                                                                 ----SGRGCLGCQGAGAAA----
                                                                                                                                                                                                                                                                                    AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA-
                                                                                                                                                                                                                                                                                                                            AAM50039 standard; Protein; 1880 AA.
                                                                                                                     GAGQGGYGGLGSQGT------
                                                                                                                                                                                                                                                                                                                                                                                                                       21-MAR-2001; 2001DE-1013781.
                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                 DE10113781-A1
                                                                                                                                                                                                                                                                                                                                                  18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                            13-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                      Synthetic.
                                                                                                                                                                                                                                                                                                                                       AAM50039;
                                                            488
                                                                        309
                                                                                              346
                                                                                                         608
                                                                                                                                664
                                                                                                                                                                  469
                                                                                                                                                                             783
                                                                                                                                                                                         523
                                                                                                                                                                                                    838
                                                                                                                                                                                                                545
                                                                                                                                                                                                                                       999
                                                                                                                                                                                                                                                  928
                                                                                                                                                                                                                                                             900
                                                                                                                                                                                                                                                                        1017
    233
                          271
                                                 293
                                                                                                                     404
                                                                                                                                                                                                                                                                                    637
                                                                                                                                                                                                                                                                                                                       AAM50039
                                                                                                                                                                                                                                                                                                                 RESULT
              셤
                                   g
                                                          g
                                                                                 g
                                                                                             ö
                                                                                                       g
                                                                                                                              g
                                                                                                                                                    유
                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                  ò
                                                                    δ
                                                                                                                   õ
                                                                                                                                                                 õ
                                                                                                                                                                          g
                                                                                                                                                                                       õ
                                                                                                                                                                                                              à
                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                              g
   ç
                          ð
                                                                                                                                          ò
                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                    ò
```

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and (iii) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Can solve they retain water solubility after long-term boiling in aqueous colutions and since they are also soluble in organic solvents but precipitated at high sait concentration, they are easily extracted and purified. The modular construction of the invention facilitates in concentration in additional peptide-encoding sequences, e.g. to simplify purification or modulate solubility. This sequence represents the synthetic N. clavipes spidroin-1 homologue SOISOISOI described in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                33;
                                                                                                                                                                                                                                                                                                    New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          72 QGGYGGLGSQGAGRGGLGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAG 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAA.----AGGAGQGGYGGLGSQ 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAA 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .---- 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 310 GGAGQGGLGGQGAGQGAGAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               127 GTSGRGGLGGQGAGAAA------AAAAAGGAGQGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----SGPGGYGP-GQQTSGRGGLGGQGAGAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels 474;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1880;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---AAAAAAGGAGQGGYGGLGSQGT-
                                                                                       (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               69.3%; Score 2419; DB 23;
52.2%; Pred. No. 6e-169;
ive 19; Mismatches 40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----SGRGGLGGQGAGAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11 GSMASGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT-
                                                                                                                                                     Guehrs K;
                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 22; Page 38-45; 88pp; German.
                                                                                                                                                     Grosse F,
09-JUN-2000; 2000DE-1028212.
24-OCT-2000; 2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                     Conrad U,
                                                                                                                                                                                                             2002-123561/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1880 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 581; Conserv
                                                                                                                                                                                                                                            N-PSDB; ABL61040
                                                                                                                                               Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Invention.
                                                                                                                                                                                                                                                                                                                                                             proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   271
```

S

```
958 AGGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGA 1016
                                                           345
                                                                                                                                                                                                                   522
                                                                                                                                                                                                                                                         544
                                                                                                                                                                                                                                                                                               565
                                                                                                                                                                                                                                                                                                                                      599
                                                                                                                                                                                                                                                                                                                                                                             636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Spidroin; spider; silk; fibre; film; membrane; wound; filter; SO1SM12.
                                                                                                 403
                                                                                                                  663
                                                                                                                                       434
                                                                                                                                                 -----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAG 468
                                                                                                                                                                                       124 AAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAG 782
                                                                                                                                                                                                                             897
547
                                                                           548 AGGGAGAAAAAAGGAGGGYGGLGSQGAGRGGQGAGAAAAAAGGAGGGGYGGLGSQGAGR 607
                                                                                                                                                                                                                                                                                                                  898 GLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGQGGLGGQGAGQGAGAAAAA 957
                                                                                                         GGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT---SGRGGLGGQGAGAAAAAAAAA
                                                                                                                                      GACQCCYGGLGSQCT------SGRGCLGGQCAGAAAA
                                                                                                                                                                                                                 469 AAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA----
                                                                                                                                                                                                                                                                                             -----SGRGGLGGQGAGAAAA------AAAAA
                                -------AAAAAGGAGQGYGGLGSQGT--------
                                                                                                                                                                                                                                                                                                                                                                           AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA-------------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              spidroin synthetic homologue SOISM12 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                           1076 GAGAAAAAAGGAGQGCYGGLGSQG-AGRGGQGAG 1108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Guehrs K;
                                                                                                                                                                                                                                                                                                                                                                                                                  ---AAAAAAGGAGGGYGGLGSQGTSGPGGYGPG 667
                                                         -----AAAAAAGGAGQGGYGGLGSQGT-----
                                                                                                                                                                                                                                                                                                                                     AGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Grosse F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAM50038 standard; Protein; 989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-MAR-2001; 2001DE-1013781.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2000DE-1028212.
2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DE10113781-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N. clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (IPKP-) IPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAM50038;
                                                                                                                                                                                                                                                                                             545
                                                         309
                                                                                                346
                                                                                                                 809
                                                                                                                                      404
                                                                                                                                                        664
                                                                                                                                                                            435
                                                                                                                                                                                                                                    783
                                                                                                                                                                                                                                                        523
                                                                                                                                                                                                                                                                                                                                     999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAM50038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                   ò
                                    a
                                                        ò
                                                                          g
                                                                                              ö
                                                                                                              a
                                                                                                                                     ò
                                                                                                                                                     g
                                                                                                                                                                       õ
                                                                                                                                                                                           ద
                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

```
34;
                                                                                                                                                        This invention describes a novel DNA sequence, encoding a synthetic sepider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (11) as adhesion surfaces for culturing cells; and (ii) as filters The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity).
                                                                                                                                                                                                                                                                                                                                                                                                            simplify
                                                                                                                                                                                                                                                                                                                                         Also they retain water solubility after long term bolling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and purified. The modular construction of the invention facilitates incorporation of additional peptide-encoding sequences, e.g. to simplification or modulate solubility. This sequence represents the synthetic N. clavipes spidroin-1 homologue SOISMIZ described in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          356 GOGAGAAAAAGGAGGGYGGIGSQGAGRGGGGAAAAAAGGAGGGCGYGGIGSQG-AGR 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----AAAAAAGGA 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA encoding synthetic spider silk protein, useful e.g. for cl
nds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11 GSMASGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GQGGYGGLGSQGT----SGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          237 GOGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LGGQGAGAAAA------AAAAAAGGAGQGGYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTS------GRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LGGQGA---GAAAAAAAAAAAGGAGQGGYGGLGSQGT------SGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels 352;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          67.9%; Score 2372; DB 23;
57.8%; Pred. No. 9.8e-166;
11ve 18; Mismatches 39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-----
                                                                                                                              Claim 22; Page 34-38; 88pp; German
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 561; Conservative
             2002-123561/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             989 AA;
                              N-PSDB; ABL61039
                                                                                                                                                                                                                                                                                                                                                                                                                                                              invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           204
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

Thu

```
The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where
                                                                                                                                                                                                                                                                                                                                                                                                                         Orb-weaving spider; silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope surgical suture; flexible tie down; electrical component; implantation; pragline protein 1; DP-1B iGmer.
                                                                                                             565
                              492
                                                  989
                                                                     531
                                                                                        695
                                                                                                                                                                                                               873
                                                                                                                                                                                                                                                      933
                                                                                                                     Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein
--- AAAAAAGGA
                                                                                                           GQGGYGGLGSQGTSGRGGLGGQGAGAAA---------AAAAA
                                                                                                                                                            ---SGRGGLGGQGAGAAA-------AAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                           AAAAAGGAGOGGYGGLGSOGT---------
                                                                                                                                                   AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                       Dragline protein 1 analogue DP-1B 16mer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          material construction, comprises provide expression cassette to a green plant -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 77-81; 93pp; English.
                                                                     GTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-
                                                                                                                                                                                                                                                                                                                                           AAU11794 standard; protein; 1617 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ы
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ္ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (DUPO ) DU PONT DE NEMOURS &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-206968P
                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                            -AGRGGGGAG 942
                                                                                                                                                                                                                                                                        TSGPGGYGPG 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200190389-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29-NOV-2001.
                                                                                                                                                                                                                                                                                                                                                                                   26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
                                                                                                                                                                                                                                                                                                                                                               AAU11794;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yang JG;
         530
                                                584
                                                                                      637
                                                                                                                                                  999
                                                                                                                                                                     755
                                                                                                                                                                                          009
                                                                                                                                                                                                             814
                                                                                                                                                                                                                                                                        658
                                                                    493
                                                                                                          532
                                                                                                                                                                                                                                621
         a
                                            8
                                                                                     셤
                                                                                                                         Q
                                                                                                                                                                     셤
                                                                                                                                                                                                        유
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                             ò
                                                                 ò
                                                                                                         ò
                                                                                                                                                 ð
                                                                                                                                                                                         à
                                                                                                                                                                                                                                ò
```

```
P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, lexible tie downs for certain electrical components, or as a biomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dragline protein (DP) I synthetic variant DPIB IGMET (1.e. 16 copies of DPIB) which is used as an SLP construct in the method of the invention.
                                                                                                                                                                                 36;
                                                                                                                                                                                                                                                                  74 QG----AGRGGQGAG----AAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGGLG
                                                                                                                                                                                                                                                                                                    136
                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         681
                                                                                                                                                                                 Gaps
                                                                                                                                                                                                         49
                                                                                                                                                                                                                   97
                                                                                                                                                                                                                                                                                                               ------AAAAAAGGAGGGYGGLGS
                                                                                                                                                                                                                                                                                                   GQGAGAAAA----------AAAAAAGGAGGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                              AAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QGTSGRGGLGGQGAGAAAAAAAAAGGGGGGYGGLGSQGT-----SGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                           AAAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAGGAGQGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA------A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------SGRGGLGGQGAGAAA-------AAAAAGGAGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---AGRGGQGAG----AAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TSGRGCLGCQGAGAAAA-------AAAAAAGGAGQGGYGGLGSQGT
                                                                                                                                                                                Indels 288;
                                                                                                                                                       Length 1617;
                                                                                                                                                                                 38;
                                                                                                                                                          DB 23;
                                                                                                                                                       Score 2294; DB 23;
Pred. No. 7.1e-160;
3; Mismatches 38;
                                                                                                                                                                               13;
                                                                                                                                                                                                      11 GSMASGRGGLGGQGAGAAA-----
                                                                                                                                                       65.7%;
                                                                                                                                                                  Best Local Similarity 61.6%
Matches 543; Conservative
                                                                                                                                 1617 AA;
                                                                                                                                 Sequence
                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            460
                                                                                                                                                                                                                                                                                                  86
                                                                                                                                                                                                                                                                                                                        26
                                                                                                                                                                                                                                                                                                                                                                                                                                            235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 909
                                                                                                                                                                                                                                                                                                                                                                                              195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               545
  8888888888888
                                                                                                                                                                                                      ð
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

us-09-490-291-4.rag

```
GQGAGAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRGG 183
                                                                                                                                                                                                                                              SQGAGQGAGAAAAAAGGAGGGGYGGLGSQG----AGRGGQGAG----AAAAAAAGGAGGGG 402
                                                                                                                                                                                                                                                                                       322 YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAAA----------A 359
                                                                                                                                                                                                                                                                                                            ----SGRGGLGGQGAGAAAA-------AAAAAAGGAGGG 446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Silk-like protein; SLP: transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flaxible tie down; electrical component; implantation; bragline protein 1; DP-1B Bmer; His tag.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           506 TSGRGGLGGQGAGAAA-----------AAAAAAGGAGQGGYGGLGSQGT
                                                                                                        -------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAAAAAA
                                                                                                                                                                    AAAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                             514 QGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           593 AGAAAA------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA
                                                                                                                                                                                                                                -----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAAGGAGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  625 GAGREGLEGQGAGAAAAAAGGAGQGGLESQGAGQAGAAAAAAGGAGQGGYGGLGSQG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                   360 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        632 GAAAAAAAAGGAGGGGYGGLGSQGTSGPGGYGP-GQQTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Oragline protein 1 analogue DP-1B/his tag 8mer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAU11797 standard; Protein; 818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200190389-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAU11797;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              684
                                                                                                        195
                                                                                                                                                                                                                              271
                                                                                                                                                                                                                                                                                                                                                                                 462
                                                                                                                                                                                                                                                                                                                                                                                                               419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               윱
                                                                                                                                                                                               g
                                                                                                                                                                                                                              ô
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                        ŏ
                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassetter. The green plant contains the expression cassetter P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, flexible tie downs for certain electrical components, or as a biomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dragline protein (DP) I synthetic variant DPLB/his tag 16mer (i.e. 16 copies of DPLB with a His stag at the C-terminus) which is used as an SLP construct in the method of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein expression cassette to a green plant -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11 GSMASGRGGLGGQGAGAAAA------AAAAAAGGAGQGGYGGLGS 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  50 QGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLG 97
                                                                                                                                                                                                                          Silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flastible tie down; electrical component; implantation; Dragline protein 1; DP-1B 16mer; His tag.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    38; Indels 288;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65.7%; Score 2294; DB 23; Length 1626; 61.6%; Pred. No. 7.1e-160; ive 13; Mismatches 38; Indels 288;
632 GAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSG 672
               Dragline protein 1 analogue DP-1B/his tag 16mer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 1; Page 87-92; 93pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ÷
                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ы
                                                                                                      AAU11798 standard; Protein; 1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ႘
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (DUPO ) DU PONT DE NEMOURS &
                                                                                                                                                                                                                                                                                                                                                                                                              24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                                            25-MAY-2000; 2000US-206968P
                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 543; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1626 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98 GOGAGAAAA---
                                                                                                                                                                                                                                                                                                       clavipes
                                                                                                                                                                                                                                                                                                                                                   WO200190389-A2.
                                                                                                                                                                  26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                 29-NOV-2001
                                                                                                                                                                                                                                                                                                                     Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                      AAU11798;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Yang JG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                       Vephila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local
                              791
                                                                          RESULT 7
ð
                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
```

505

624

- AAAAAAGGAGOG

683

624

```
in a green plant by providing a green plant containing an Street Parts of cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing sliks and silk-like proteins in green plants. The silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in flabrics or in material construction, such as rope, surgical sutures, flexible tie downs for certain electrical components, or as a blomaterial for implantation. The method allows for more cost effective
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Orb-weaving spider; silk-like protein; SLP; transgenic plant; prometer; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation; bragline protein 1; DP-1B Smer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to a method of producing silk-like proteins (SLP)
                                                                                                                                                                                                                                                                                                                                                              684 ---AGRGGQGAG----AAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQG 736
                                                                                                                                                                                                                                                                                                                                                                                                                                                            737 AGAAAAAAGGAGGGLGSQGAGQGAGAAAAAGGAGQGGYGGLGSQG----AGRGGQGA 792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein expression cassette to a green plant -
                                                         GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQ
                                                                                                                                                                                                                        ----- AAAAAAGGAGQGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                 SGRGGLGGQGAQAAAAAAAAAGGAGQGGYGGLGSQCT-----SGRGGLGGQG
                                                                                                                                                                                                                                                                  625 GAGRGGLGGQGAGAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dragline protein 1 analogue DP-1B 8mer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Page 74-76; 93pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Е
П
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ä.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           812
                             - SGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAU11793 standard; protein; 809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (DUPO ) DU PONT DE NEMOURS & CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GAAAAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     G----AAAAAAGGAGGGGGGGGGGG
                                                                                                                                                                                                                     506 TSGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25-MAY-2000; 2000US-206968P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                AGAAAA - - - - - - - - - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nephila clavipes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200190389-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29-NOV-2001,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Yang JG;
                                                                        514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        793
                                                                                                                                                                                                                                                                                                                                                                                                             593
                                                                                                                       447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                     a
                                                                                                                                                                 g
                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                              유
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                            in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing silks and silk like proteins in green plants. The silks and silk-like proteins may be used in flathics or in material construction, such as rope, surgical sutures, blomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dragline protein (DP) is synthetic variant DPLB/his tag 8mer (i.e. 8 copies of DPLB with a His tag at the C-terminus) which is used as an SLP construct in the method of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   34;
                                                                                                                                                                                                                                                                                                                                                                                                        The invention relates to a method of producing silk-like proteins (SLP)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QGAGAAAAAAAAAAAAAGGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAAA- 194
                                                                                                                                                                                                                                                           Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein expression cassette to a green plant -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QCTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLG 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------AAAAAGGAGGGYGGLGS 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  239 AAGGAGGGLGSQGAGQGAGAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGGGYGGLGSQGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAAAGGAGOGGYGGLGSQGT-----SGRGGLGGOGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 818;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64.5%; Score 2254; DB 23;
61.7%; Pred. No. 3.4e-157;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 3.4e-157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11; Mismatches
                                                                                                                    Е І.
                                                                                                                                                                                                                                                                                                                                                              Example 1; Page 83-86; 93pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11 GSMASGRGGLGGQGAGAAA-----
                                                                                                                  (DUPO ) DU PONT DE NEMOURS & CO
                                                                   25-MAY-2000; 2000US-206968P
                      24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 533; Conservative
                                                                                                                                                                                                              WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    818 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                 Yang JG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   462
$\times \times \
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
```

Š

AAM50042 standard; Protein; 630

AAM50042

AAM50042;

```
34;
                                                                                                                                                       98 GQGAGAAAA------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 136
                                                                                                                                                                                      288
                                                                                                                                                                                                                                                                 359
                                                                                                                                                                                                                                                                                                                                                                                        446
                                                                                                                                                                                                                                                                                                                                                                                                                      505
                                                                                                                                                                                                                                                                                                                                                                                                                                      622
                                                                                                                               234
                                                                                                                                                                                                                                                                                        321
                                                                                                                                                                                                                                                                                               511
                                                                                                                                                                                                                                                                                                                                                                                                                                                       544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                      571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                 as
                                                                                        --- AAAAAAGGAGQGGYGGLGS 49
                                                                                                                       QCTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLG 97
                                                                                                                                                                is used
                                                                                                                                                                                                                        237 AAGGAGQGGLGSQGAGQGAGAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AA
                                                                                                                                                                                                                                                                                                                                                                 -----SGRGGLGGQGAGAAA------AAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                               AAAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                        -----AAAAAAGAGGYGGIGSQGTSGRGLIGGQGAGAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----AAAAAAGGAGOGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT------SGRGGLGGQG
  sonrces
                                                                        286;
           (PP)
                                                        Length 809;
production of silk not obtained from natural or microbial The present sequence is the silk protein Dragline protein synthetic variant DPIB Wher (i.e. 8 copies of DPIB) which an SLP construct in the method of the invention.
                                                                                                                                                                                                                                                                                                                        YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA--------
                                                                        Indels
                                                        64.4%; Score 2250; DB 23;
61.6%; Pred. No. 6.7e-157;
ive 11; Mismatches 34;
                                                                                        GSMASGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              506 TSGRGGLGGQGAGAAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GAAAAAAAAAGGGGGGGGGGGG 654
                                                               Local Similarity 61.6
les 532; Conservative
                                        809 AA;
                                         Sequence
                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                       360 7
                                                                 Best Loc
Matches
                                                                                        11
                                                                                                                       20
                                                                                                                                                                                                                       195
                                                                                                                                                                                                                                                                                       271
                                                                                                                                                                                                                                                                                                                       322
                                                                                                                                                                                                                                                                                                                                                                                      419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      791
 88888
                                                                                        à
                                                                                                      셤
                                                                                                                       ò
                                                                                                                                     g
                                                                                                                                                      ò
                                                                                                                                                                     g
                                                                                                                                                                                       à
                                                                                                                                                                                                    g
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

RESULT 10

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonuclotide encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or support or cover artificial organs; (11) as adhesion surfaces for culturing cells; and artificial organs; (11) as adhesion surfaces for culturing cells; and cills as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Can be prepared at high salt concentration, they are easily extracted and precipitated at high salt concentration, they are easily extracted and incorporation of additional peptide-encoding sequences, e.g. to simplify the synthetic N. clavibes spidroin-1 homologue SOI described in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       useful e.g. for closing units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               71 AAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGGYGGLGSQGTSG 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAG 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11 GSMASGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels 174;
                                                                                                                               Spidroin; spider; silk; fibre; film; membrane; wound; filter;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length
                                                                                               N. clavipes spidroin synthetic homologue SO1 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 2139; DB 23;
Pred. No. 6.9e-149;
9; Mismatches 42;
                                                                                                                                                                                                                                                                                                                                                    (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New DNA encoding synthetic spider silk protein, wounds, comprises modules that encode repeating
                                                                                                                                                                                                                                                                                                                                                                                       ×
                                                                                                                                                                                                                                                                                                                                                                                       Guehrs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 22; Page 48-50; 88pp; German
                                                                                                                                                                                                                                                                                                                                                                                       Grosse F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61.2%;
68.7%;
                                                                                                                                                                                                                                                                  21-MAR-2001; 2001DE-1013781
                                                                                                                                                                                                                                                                                                 09-JUN-2000; 2000DE-1028212
24-OCT-2000; 2000DE-1053478
                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 493; Conservative
                                                                                                                                                                                                                                                                                                                                                                                     Scheller J, Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2002-123561/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               630 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; ABL61043
                                                                                                                                                                                              DE10113781-A1.
                                                              18-SEP-2002
                                                                                                                                                                                                                                  13-DEC-2001.
                                                                                                                                                                Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  131
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
AAAAAAAAA ------AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGA 241
                                   301
                                                                   460
                                                                                                                                                  571
                                                     254
                                                                                     255 GAG----AAAAAAGGAGGGYGGLGGQG-AGQGGYGGLGSQGAGRGGLGGQGAGAAAAA 309
                                                                                                       GLGGQGAGAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAA 520
                                                                                                                                                                                                                       ------AAAAAAGGAGGG 611
                                                                                                                                                                                                                                                                                       YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT--SGPGGYGPG 667
                                            158 AAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQG-AGRGGLGGQGAG----AAAAAAAGGA
                                  GOGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQ
                                                                                                                                         AAGGAGOGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG
                                                                                                                                                                                                            Spidroin; spider; silk; fibre; film; membrane; wound; filter; SO1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Synthetic spidroin homologue SO1"
                                                                                                                                                                                                                                                                                                                                                                                              N. clavipes spidroin synthetic homologue SO1 protein #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note= "c-myc-tag"
673..676
/note= "ER retention signal"
                                                                                                                                                                                                                                               GGYGGLGSQGTSGRGGLGGQGAGAAA-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Guehrs K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           1..28
/label⇔ LeB4_signal_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                              Ä.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Grosse F,
                                                                                                                                                                                                                                                                                                                                           AAM50047 standard; Protein; 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000DE-1028212.
2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-MAR-2001; 2001DE-1013781
                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note= "S
660..672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29..659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DE10113781-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09-JUN-2000;
24-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                              18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic.
                                                                                                                                                                                                                                                                                                                                                              AAM50047;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Key
Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Domain
191
                                  242
                                                                                                       359
                                                                                                                                         401
                                                                                                                                                         365
                                                                                                                                                                          461
                                                                                                                                                                                                                                                                                612
                                                                                                                                                                                                                                                                                                267
                                                                                                                                                                                                                                                                                                                                    AAM50047
õ
                 g
                                                셤
                                                                   ö
                                                                                    g
                                                                                                     à
                                                                                                                    g
                                                                                                                                        ð
                                                                                                                                                    g
                                                                                                                                                                      à
                                                                                                                                                                                        g
                                                                                                                                                                                                            à
                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                             à
```

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein. Comprising modules, each comprising a group of sequentially arranged oilgonucleidies, each oilgonucleidie encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and (iii) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Also they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and purified. The modular construction of the invention facilitates construction of additional peptide-encoding sequences, e.g. to simplify purification or modulate solubility. This sequence represents a construct composed of the Leba signal peptide-encoding sequence represents a construct composed of the Leba signal peptide-encoding sequence represents a construct composed of the Leba signal peptide. N. clavipes spidroin.1 synthetic homologue SOl, a c-Myc-tag and an endoplasmic reticulum (ER)-retention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31;
                                                   encoding synthetic spider silk protein, useful e.g. for closing comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RGGLGGGGAAAAAAAAAAAGAGGGYGGCGSGGTSGPGGYGPGQQTSGRGGLGGGGGA 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSG 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    144 -AGRGGGGAG----AAAAAAGGAGGGYGGLGSQG-------AGRGGLGGQGAG 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                302 GAGAAAAAAAAAAGGAGGGYGGLGSQCTSGPGGYGP-GQQTSGRGGLGGQGAGAAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          521 AAAAAAAAGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAA-----AGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLGGQGAGAAAAAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            242 GQGGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQCTSGRGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 174;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 676;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. 7.3e-149;
9; Mismatches 42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 2139;
                                                 New DNA encoding synthetic spider silk
                                                                                                                        Example 1; Fig 10A; 88pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61.2%;
68.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
              WPI; 2002-123561/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best_Local Similarity
Matches 493; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   676 AA;
                                                                                      proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
Orb-weaving spider; spidroin 1; silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation.
                                                                                                                                                                                   AGGAGOGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAA.....AGGAGQGGYGGLG 250
                                                                                         SQGTSGRGGLGGQGAGAAAAAA-------AAAAGGAGGGYGGLGS 289
                                                                                                                                                                                                                                                 QGTSGRGGLGGQGAGAAAAAAAAGGGGGGGGGGGGGGGGTSGPGGYGP-GQOTSGRGGL 348
                                                                                                                                                                                                                                                                                                                                      GGQGAGAAAA------AAAAAAAGGAGQGGYGGLGSQGTSGRGCLGGQG 390
                                                                                                                                                                                                                                                                                                                                                             39]. AGAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGGYGG 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                LGSOGTSGRGGLGGQGAQAAAAAAAAGGAGGGYGGLGSQGT-----SGP 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA.----AGGAGQGGYGGLGSQGTSGRGGL 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               591 QGAGAAAAAAAAGGAGGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAGGGGGY 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    532 GGGGAGAAAAAAGGAGGGLGGQGAGQGAGAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGQGAGAAAA----------AAAAAAGGAGQGGYGGIGSQGTSGRGGIGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              natural silk protein Spidroin 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAU11781 standard; protein; 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (DUPO ) DU PONT DE NEMOURS & CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAY-2000; 2000US-206968P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2002-106209/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nephila clavipes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200190389-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       651 GGLGS 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CGVGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Yang JG;
                                                                                                              143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     371
                                                                                                                                                          251
                                                                                                                                                                                                                                               290
                                                                                                                                                                                                                                                                                                                                      349
                                                                                                                                                                                                                                                                                                                                                                             312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                à
                                                                                                  g
                                                                                                                                                       ò
                                                                                                                                                                                              g
                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cosmetic or dermatological composition containing spider silk protein, for hair or skin care, in make-up or sunscreens
--AAAAAGGAGGGYGGLGSQG----AGRGGQGAGAAAAAAGGAGGGYGGLGGQGAGQ 535
                                                                572 GGYGGLGSQGTSGRGCLGGQGAGAAAA--------AAAAAAGGAGQGG 611
                                                                                                            GGYGGLGSQG-AGRGGLGGQGAGAAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGG 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present sequence represents the natural spider silk protein spidroine major 1. The protein improves the moisturizing/softening action of the compositions. The protein, and its fragments are used in cosmetic or dermatological compositions. These compositions have use as hair or skin care products; and make-up or sunscreens. As the protein is a good, persistent film-formers on the skin of low surface density, it can be used for delivery of active agents that are generally difficult to administer, e.g. vitamins, hormones, moisturizers or agents for treating disorders of the skin and hair.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       612 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT--SGPGGYGPG 667
                                                                                                                                                                              23 QGAGAAAAAAAAAAGGAGQGGYGGLGSQGT--SGRGGLGGQGAG-AAAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61.1%; Score 2135; DB 20; Length 651; 68.3%; Pred. No. 1.4e-148; 1ve 15; Mismatches 39; Indels 176;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Spider silk protein; spidroine major 1; cosmetic; make-up; dermatological compositions; hair care; skin care; sunscreen; hormone; moisturizer; skin disorder; skin disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                              Spider silk protein spidroine major 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Arrandeau JP;
                                                                                                                                                                                                                                                                                                             AAY40097 standard; protein; 651 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 3; Fig 1; 32pp; French
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98FR-0001614.
                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 68.3
Matches 495; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Philippe M, Garson JC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1999-510729/43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (OREA ) L'OREAL SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            651 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FR2774588-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                19-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-AUG-1999
                                                                                                                                                                                                                                                                                                                                                      AAY40097;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                         536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     80
                                                                                                                                                                                                                                                               RESULT 1
                                                                                                         셤
                                                                                                                                                     ŏ
                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                             THE STATE OF THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

us-09-490-291-4.rag

```
The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a proteins. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, blomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is orb-weaving spider spidoill, a natural silk protein whose sequence is used as the basis for the SLP proteins used in the method of the invention.
                                                                                                                                                                                                                                                                                                                                                                              34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  140 GAAAAAAAAAAGGAGGGYGGLGSQGTSGPGQQTSGRGGLGGQGAGAAAAAAAA 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        G---AAAAAAAGGAGGGYGGLGNQG------AGRGGQG-----AAAAA 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AGGAGQGGYGGLGSQGTSGRGGLGCQGAGAAAAAAAA------AGGAGQGGYGGLG 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   531
                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             587
                 useful in fabrics and
                                                                                                                                                                                                                                                                                                                                                                                                      23 QGAGAAAAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGAG-AAAAAAAAAAGGAGQ 79
                                                                                                                                                                                                                                                                                                                                                                                                                                  1 QGAG----AAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGGQGAGQGAGAAAAAAAAAGGAGQ 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGQGAGAAAA--------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SQGTSGRGGLGGQGAGAAAAAA........AAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AGAAAAAAAAAAGGGGYGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGYGGLGSQGAGR---GGQGAGAAAANVGAGQEGIRQQGAGQGGYGGLGSQG-SGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GGQGAGAAAA--------AAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGQGAGAAAAAAGGAGGGLGGQGAGQCAGAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGT-----SGP
                               material construction, comprises providing a silk-like protein expression cassette to a green plant -
                                                                                                                                                                                                                                                                                                                                                                             Indels 176;
                                                                                                                                                                                                                                                                                                                                                 Length 651;
                                                                                                                                                                                                                                                                                                                                              DB 23;
1.4e-148;
nes 39;
                 plant,
                                                                                                                                                                                                                                                                                                                                              5; Score 2135; D6; Pred. No. 1.4e15; Mismatches
                 Producing silk-like proteins in a green
                                                                       Disclosure; Page 69-71; 93pp; English
                                                                                                                                                                                                                                                                                                                                              61.1%;
68.3%;
                                                                                                                                                                                                                                                                                                                                                                      Matches 495; Conservative
                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                651 AA;
                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      591
Q
                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
```

```
34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The present sequence represents a spider silk protein from the present invention. Spider silk proteins, and peptide fragments within the proteins, can be produced and purified independently and can then be mixed and made into fibres that have higher tensile strengths and elasticity than naturally occurring fibres. The fibres can be used in mixed composites. The invention allows the two naturally occurring Nephila clavibes silk proteins to be produced independently so that they can later be combined to form silk fibres of high tensile strength
                                                                                                                                                                                                         dragline;
588 QGAG----AAAAAGGAGGGYGGLGGQGV-GRGGLGGQGAG-----AAAAGGAGQGGY 636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      200 AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAA-----AGGAGQGGYGGLG 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 QGAG----AAAAAAGGAGGGYGGLGGQGAGQGGYGGLGGQGAGGAGAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       80 GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 176;
                                                                                                                                                                                                         tandem repeat; flbre;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 718;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Recombinant spider silk proteins - useful for making fibres
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61.1%; Score 2135; DB 19;
68.3%; Pred. No. 1.5e-148;
1ve 15; Mismatches 39;
                                                                                                                                                                                                     Nephila clavipes; silk protein;
tensile strength; elasticity.
                                                                                                                                                                              Nephila clavipes spider silk protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 1; Column 29-34; 68pp; English.
                                                                                                            ¥
                                                                                                         standard; Protein; 718
                                                                                                                                                                                                                                                                                                                                                 90US-0511792.
94US-0317844.
95US-0425069.
                                                                                                                                                                                                                                                                                                                                                                                                                         Ξ
                                                                                                                                                                                                                                                                                                               95US-0425069
                                                                                                                                                                                                                                                                                                                                      91US-0684819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                        Lewis RV,
                                                                                                                                                                                                                                                                                                                                                                                                (UYWY-) UNIV WYOMING
                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1998-270437/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    718 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAV23249
                                                                                                                                                                                                                                       Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           and elasticity.
                        GGLGS 655
                                                                                                                                                                                                                                                                                                              19-APR-1995;
                                                                                                                                                                                                                                                                                                                                     15-APR-1991;
                                                                                                                                                       06-JUL-1998
                                                                                                                                                                                                                                                               US5728810-A.
                                                                                                                                                                                                                                                                                                                                                  20-APR-1990;
04-0CT-1994;
                                     ||:||
GGVGS
                                                                                                                                                                                                                                                                                      17-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                       19-APR-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 495;
                                                                                                                                                                                                                                                                                                                                                                                                                        Hinman MB,
                                                                                                         AAW53346
                                                                                                                                AAW53346;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                       Spider;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local
                                              637
                       651
                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       140
                                                                                               AAW53346
                                                                                                                     셤
                       ò
                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        650
```

65pp; English

```
Fig 6A-D;
                                                                                     tches 495;
                                                                                                                                                                                                                                                                                                                                                                                                     CGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                 GGVGS
                                                              Seguence
Claim 1;
                                                                         Query Match
                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                   Search comp
Job time :
                                                                                                 23
                                                                                                                          80
                                                                                                                                     57
                                                                                                                                                                                                                                                  34.9
                                                                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                       476
                                                                                                                                                  140
                                                                                                                                                              108
                                                                                                                                                                                                  251
                                                                                                                                                                                                              202
                                                                                                                                                                                                                          290
                                                                                                                                                                                                                                       261
                                                                                                                                                                                                                                                                           391
                                                                                                                                                                                                                                                                                       371
                                                                                                                                                                                                                                                                                                    451
                                                                                                                                                                                                                                                                                                                            498
                                                                                                                                                                                                                                                                                                                                                    551
                                                                                                                                                                                                                                                                                                                                                                            591
                                                                                                                                                                                                                                                                                                                                                                                                    651
 8 x 5 5 5 5 5 5 5 5 x 8
                                                                                                                                     g
                                                                                                                                                            g
                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                  ð
                                                                                                            g
                                                                                                                          ö
                                                                                                                                                  à
                                                                                                                                                                          ò
                                                                                                                                                                                     d
                                                                                                                                                                                                ò
                                                                                                                                                                                                            g
                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                               g
                  289
                             260
                                          348
                                                                  390
                                                                              370
                                                                                          450
                                                                                                       419
                                                                                                                               475
                                                                                                                                            550
                                                                                                                                                       531
                                                                                                                                                                   590
                                                                                                                                                                                            650
                                                                                                                                                                                                       989
                                                                                                                   497
                                                                                                                                                                               587
                                                      311
      201
                                                                                                                                                                                                                                                                                                                                                                                                                                         ħ
                                                                                                                                                                                                                                                                                                                protein; dragline silk protein; major ampullate gland;
GGQGAGAAAA-------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                   LGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGP
                                                                                                                                                                                                  QGTSGRGLGGQGAGAAAAAAAAAAGGAGGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                          AGAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG
                                                                                                                                            GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                   GCQCACAAAA---------AAAAAAGGGGGGGGCIGSQCTSGRGCLGG
                                                                                                                                                                               532 GGGGAGAAAAAAGGAGGGLGGQGAGQGAGAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                            QGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAGGAGQGGY
                  ------AAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                         and useful
                                                                                                                                                                                                                                                                                                                                                                                                                                        Isolated DNA, vector and transformed cell encoding for the production of spider silk protein
                                                                                                                                                                                                                                                               Ž
                                                                                                                                                                                                                                                                                                    silk protein 1.
                  SOGTSGRGGLGGQGAGAAAAAA - - - - -
                                                                                                                                                                                                                                                              AAY59070 standard; Protein; 718
                                                                                                                                                                                                                                                                                                                                                                                                          ₹,
                                                                                                                                                                                                                                                                                                                                                                           91US-0684819.
                                                                                                                                                                                                                                                                                                                                                                94US-0317844
                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                          Hinman MB, Xu M, Lewis
                                                                                                                                                                                                                                                                                                                                                                                              (UYWY-) UNIV WYOMING.
                                                                                                                                                                                                                                                                                                                                                                                                                      2000-061225/05.
                                                                                                                                                                                                                                                                                                    spider
                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-061225/
N-PSDB; AAZ38195
                                                                                                                                                                                                                                                                                                                           Nephila clavipes
                                                                                                                                                                                                                   GGLGS 655
                                                                                                                                                                                                                               641
                                                                                                                                                                                                                                                                                                                                                               04-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                  20-APR-1990;
                                                                                                                                                                                                                                                                                                               Spider silk
                                                                                                                                                                                                                                                                                                                                       US5989894-A.
                                                                                                                                                                                                                                                                                                                                                                            15-APR-1991;
                                                                                                                                                                                                                                                                                       08-MAR-2000
                                                                                                                                                                                                                                                                                                   N. clavipes
                                                                                                                                                                                                                                                                                                                                                    23-NOV-1999
                                                                                                                                                                                                                           ||:||
||:||
                                                                                                                                                                                                                               637
                                                                  349
                                                                                                                                                       476
                                                                                                                                                                                                       588
                              202
                                          290
                                                                                          391
                                                                                                      371
                                                                                                                              420
                                                                                                                                                                                           591
                                                                                                                                                                                                                   651
      143
                 251
                                                      261
                                                                                                                   451
                                                                                                                                           498
                                                                                                                                                                   551
                                                                                                                                                                                                                                                  15
                                                                                                                                                                                                                                                  RESULT 15
                 à
                            g
                                                   Q
                                                                            a
                                                                                       ŏ
                                                                                                   g
                                                                                                                ò
                                                                                                                            qq
                                                                                                                                          õ
                                                                                                                                                   ద
                                                                                                                                                                  ò
                                                                                                                                                                            g
                                                                                                                                                                                         ò
                                                                                                                                                                                                    Q
                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                               δ
                                                              ò
```

```
34;
The invention provides isolated cDNA molecules coding for spider silk proteins. The spider silk proteins are characterized by repeating alpha and beta regions and optional variable regions. The DNA sequences are useful in the production of spider silk protein by recombinant DNA techniques. The recombinant spider silk proteins may be used for the production of fibers. The present sequence represents the spider silk protein 1, derived from the major ampullate gland of Nephila clavibes.
                                                                                                                                                                                                                                                                                                                   139
                                                                                                                                                                                                                                                                                                                                                                                         199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    636
                                                                                                                                                                                                                                                                                                                                                     107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            201
                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                          56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QGAGAAAAAAAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGAG-AAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                   GGYGGLGSQG----AGRGGQGAG----AAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                    200 AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA....-AGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGQGAGAAAA------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SQGTSGRGGLGGQGAGAAAAAA-------AAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGAAAAAAAAAAGGGGYGGLGSQGTSCRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGQGAGAAAA-------------AAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 532 GGQGAGAAAAAAGGAGGGLGGQGAGQGAGAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QGAGAAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGY
                                                                                                                                                                          61.1%; Score 2135; DB 21; Length 718; 68.3%; Pred. No. 1.5e-148; 1ve 15; Mismatches 39; Indels 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            December 18, 2002, 16:17:12
                                                                                                                                                                                                              Conservative
                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            completed: Decemine: 51.4697 secs
                                                                                                                                          718 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        641
```

This Page Blank (uspto)

Н

```
December 18, 2002, 16:14:14; Search time 15.6792 Seconds (without alignments) 1277.936 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                             US-09-490-291-4
3493
1 MRGSHHHHHGSMASGRGGL.....GGYGPGQQTSGIRRPAAKLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                          of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                 262574 seqs, 29422922 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                             protein search, using sw model
                                                                                                                                                                                                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          length: 0
length: 2000000000
                                                                                                                                                                                                                               Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            seq
                                                                                                                                                                                                                                                                                                                             Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                               Total number
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB s
Maximum DB s
                                                                                         OM protein
                                                                                                                                                                                                                                                                                                                                                                                                   Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database
                                                                                                                                         Run on:
```

1, Appli 22, Appli 33, Appli 21, Appli 22, Appli 8, Appli 6, Appli 6, Appli 61, Appli Sequence 19, Appl Sequence 1, A Sequence 2, A Sequence 2, A Sequence 2, A Sequence 4, A Sequence 6, A Sequence 6, A Sequence 6, A Sequence 10, Sequence 10, Sequence 11, Sequence 1 Description US-08-556-978B-19 US-09-247-806-1 US-08-1425-069-2 US-08-147-844B-2 US-09-034-177-3 US-09-556-978B-21 US-09-247-806-8 US-08-556-978B-63 US-08-256-978B-63 US-08-247-806-6 US-08-256-978B-61 US-08-247-806-10 US-08-247-806-10 US-08-247-806-10 US-08-247-806-10 US-08-247-806-10 US-08-247-806-10 US-08-26-978B-61 US-08-26-978B-64 US-08-26-978B-64 SUMMARIES В Query Match Length

2135 2135 2135 2135 2036 2076 2076 1972.5 1971.5 1961.5 1405

Result Š. 5-08-475-411A-31 5-08-478-029A-31 5-09-444-791A-64

1331.5 1233.5 1233.5 1233.5 1233.5 1233.5 1233.5 1233.5 1176.5 1176.5

US-08-707-237A-54 US-08-806-029-10

Appl Appl Appli Appli Appli Appl Appl Ap		34;
11.5999961114		Gaps .GQ 79
sedneuce Sed		ength 651; Indels 176; G AAAAAAAAAGGAGQ
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		Length 6 Indels -AAAAAA
	COMPANY	.⊒ ·
S.08-477-509B-83 S-08-482-085B-83 S-09-444-791A-83 S-08-464-038A-3 S-08-425-069-4 S-08-425-069-4 S-08-475-41A-41 S-08-478-029A-41 S-08-478-029A-41 S-08-478-029A-41 S-08-478-029B-89 S-08-478-029B-89 S-08-477-509B-89 S-08-478-028B-89 S-08-478-085B-89 S-08-478-085B-89 S-08-478-085B-89 S-08-478-085B-89 S-08-478-085B-89 S-08-478-085B-89 S-08-478-085B-89 S-08-478-085B-89 S-08-175-155-54		135; DB 4; o. 7.2e-151; atches 39; -SGRGGLGGQGAG
477 - 50 4844 - 79 4844 - 79 4	ALIGNMENT 78B F. MBINANTLY R ANALOGS NEMOURS A WINDOWS 6,978B 6,978B	NŽĖ
US - 08 - 08 - 08 - 08 - 08 - 08 - 08 - 0	ALIGNMENTS US/08556978B K, STEPHEN F. NOVEL RECOMBINANTLY PR SPIDER SILK ANALOGS 107 S: S: TO WITH AN OUR STREET ATES OF AMERICA MICROSOFT WINDOWS 95 MICROSOFT WINDOWS 95 ATES OF AMERICA MICROSOFT WINDOWS 95 ATES OF AMERICA MICROSOFT WINDOWS 95 ATES OF AMERICA MOND FOR WINDOWS 95 ATA: US/08/556,978B 5 ATA: 193,692 ATANETHY 15,193 ATION: 15,193 ATION: 164 ATANETHY 170 164 170 181 164 181 186 OWN 19: 105:	SO Mi
46464444444444444444444444444444444444	/0855 STEPH EL RE DER S 7 7 3.5 S OF S OF S OF S OF S OF S OF S OF S OF	18; 38; 15 3GYGC
11101 11101 11011 11011 1338 1338 1038 10	ALIGNMEN. 101-1 101-556-978B-19 equence 19, Application US/08556978B equence 19, Application US/08556978B estent No. 6268169 ESTREAT. INFORMATION: TITLE OF INVENTION: TITLE OF INVENTION: CORRESPONDENCE: 107 CORRESPONDENCE: 107 CORRESPONDENCE: 1 DU PONDE SILK ANALOGI NUMBER OF SEQUENCES: 107 CORRESPONDENCE: 1 DU PONDES STREET: 1007 MARKET STREET CITY: WILMINTON STREET: 1007 MARKET STREET COMPTRY: UNITED STATES OF AMERICA ZIP: 19898 COMPTRY: UNITED STATES OF AMERICA ZIP: 19808 COMPTRY: UNITED STATES OF AMERICA STREET: 1007 MARKET STREET COMPTRY: WILMINTON STREET: 1993 MEDIUM TYPE: DISKETTE, 3.50 INCH COMPUTER: IBM PC COMPATIBLE OMPOUTER: IBM PC COMPATIBLE COMPUTER: IBM PC COMPATIBLE OMPOUTER: IBM PC COMPATIBLE COMPUTER: IBM PC COMPATIBLE COMPUTER: IBM PC COMPATIBLE COMPUTER: IBM PC COMPATIBLE OMPOUTER: INTROGENTY WINDOWS SOFTWARE: MICROSOFT WINDOWS CLASSIFICATION NUMBER: 08/077,600 FILING DATE: JUNE 15, 1993 ATTORNEY AGENT INFORMATION: TELEPHONE: 302-892-8112 TENEDRATICATION NUMBER: 33,692 RATRANDEDNESS: UNINCOMPANINCE TENEDRATICATION NUMBER: 33,692 TENEDRATICATION NUMBER: 33,692 ATTORNEY NUMBER: 33,692 TENEDRATICATION NUMBER: 33,692 ATTORNEY NUMBER: 33,692 TENEDRATICATION NUMBER: 33,692 TENEDRATICATION	61.1%; Sc al Similarity 68.3%; Pr 495; Conservative 15; QGAGAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGG
66699999999999999999999999999999999999	T. 1 -556-978B-19 uence 19, Application Unence 19, Application Uneral Information Of 268B19 APPLICANT: FAHNESTOCK, TITLE OF INVENTION: NO TITLE OF INVENTION: NO ADDRESSED DELAWARE OF STATE: D107 MARKET CITY: WILMINTON STATE: D1809BBESSED OF THE D180BBESSED OF THE D180B	h Similarity 68 95; Conservativ AGAAAAAAAGGAG
76.5 115.7 113.2 113.2 113.2 113.2 112.9 112.9 112.7 10.7 10.7 10.7 10.7 10.7 10.7 10.7 10	ULT 1 08-556-978B-19 equence 19, Applica green No. 6268169 GREEAL INFORMATION APPLICANT: FAHNE TITLE OF INVERTION TITLE OF INVERTION TITLE OF SEQUENC CORRESPONDENCE AD CORRESPONDENCE AD STREET: 1007 M CITY: WILMINTO STREET: DELAWAR COUNTRY: WILMINTO STREET: DELAWAR COMPUTER READABLE MEDIUM TYPE: DE COMPUTER READABLE TELECOMUNICATION NUM FILING DATE: J ATTORNEY/AGENT IN NAME: FLOYD, L REGISTRATION NUM FILING DATE: J ATTORNEY/AGENT IN NAME: FLOYD, L REGISTRATION NUM FILING DATE: J ATTORNEY/AGENT IN NAME: FLOYD, L REGISTRATION NUM FILEPHONE: 302-7 INPORMATION FOR SEG SEQUENCE CHARACTE LENGTH: 651 am TYPE: amino ac STRANDEDNESS: TOPOLOGY: UNKN MOLECULE TYPE: p	tch al Simi: 495; (OGAGAAA
110031	1 10ce 19, 14 No. 62 17 LE OF 1 17 LE OF 1 18 MBER OF 1 19 COMPUTER 1 20 CONTRY: 0 21 P: 19 COMPUTER 1 21 LI LOR DE 1 21 LE OF 1 21 LE OF 1 22 LE OF 1 23 LE OF 1 24 PELICATER 1 25 LE OF 1 26 LE OF 1 27 LE OF 1 27 LE OF 1 28 COMPUTER 1 28 COMPUTER 1 28 COMPUTER 1 29 COMPUTER 1 20 COMPUTER 1 20 COMPUTER 1 21 LE OF 1	Matc Local es 4
	S-O8 Seq Pat Pat GG GS-O8	Oue Bes Mat
	# D ************************************	δ

US-09-247-806-1

```
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: AFRADUBAU, Jean-Claude
APPLICANT: AFRADUBAU, Jean-Pletre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
SEARLIER APPLICATION NUMBER: FR 98/01614
ERRLIER FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOSTWARE: PATENTIN VET: 2.1
                                                                                                                                                                                                       348
                                                                                                                                                                                                                                                   390
                                                                                                                                                                                                                                                                          370
                                                                                                                                                                                                                                                                                              590
                                                                                                                                                                                                                                                                                                                                                                                                                                                       587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               650
                                                                              260
                                                                                                                                      201
                                                                                                                                                            289
                                                                                                                                                                                                                             311
                                                                                                                                                                                                                                                                                                                                                     550
                                                                                                                                                                                                                                                                                                                                                                                                   636
                                  26
                                                                                                                                                                                                                                                               QGTSGRGGLGGQGAGAAAAAAAAAAGGAGGAGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                 GGQGAGAAAA---------AAAAAAGGAGGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                               QGAGAAAAAAAAAAGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGGGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QGAG----AAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGGQGAGQGAGAAAAAAAGGAGQ
                      GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA
                                                                 AGGAGQGGYGCLGSQGTSGRGCLGGQGAGAAAAAAAAA...---AGGAGQGGYGGLG
                                                                                                                                                          SQGTSGRGGLGGQGAGAAAAAA-----------AAAAGGAGQGGYGGLGS
                                                                                                                                                                      GGQGAGAAAA-------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                      GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA.----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 1, Application US/09247806
Patent No. 6280747
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT
ORGANISM: Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ||:||
|GGVGS 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-247-806-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQ ID NO 1
                                                                                                              200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   588
                      80
                                                                                         108
                                                                                                                                                          251
                                                                                                                                                                                202
                                                                                                                                                                                                       290
                                                                                                                                                                                                                            261
                                                                                                                                                                                                                                                  349
                                                                                                                                                                                                                                                                                                                   371
                                                                                                                                                                                                                                                                                                                                           451
                                                                                                                                                                                                                                                                                                                                                                                      498
                                                                                                                                                                                                                                                                                                                                                                                                                                  551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              637
                                                                                                                                                                                                                                                                                               391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         651
                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                         셤
셤
                      ò
                                           셤
                                                                   à
                                                                                      Ω
                                                                                                             ç
                                                                                                                                q
                                                                                                                                                      ò
                                                                                                                                                                             q
                                                                                                                                                                                                     ò
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                              Ω
                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 2, Application US/08425069
Patent No. 5728810
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: ALWING
APPLICANT: ALMING
APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
                                                                           139
                                                                                             107
                                                                                                                199
                                                                                                                                   142
                                                                                                                                                     250
                                                                                                                                                                        201
                                                                                                                                                                                           289
                                                                                                                                                                                                              260
                                                                                                                                                                                                                                 348
                                                                                                                                                                                                                                                   311
                                                                                                                                                                                                                                                                      390
                                                                                                                                                                                                                                                                                         370
                                                                                                                                                                                                                                                                                                            450
                                                                                                                                                                                                                                                                                                                             419
                                                                                                                                                                                                                                                                                                                                                497
                                                                                                                                                                                                                                                                                                                                                                  475
                                                                                                                                                                                                                                                                                                                                                                                      550
                                                                                                                                                                                                                                                                                                                                                                                                        531
                                                                                                                                                                                                                                                                                                                                                                                                                           590
                                                                                                                                                                                                                                                                                                                                                                                                                                             587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   989
                   Gaps
                                     79
                                                       26
                                                                                    CGYCGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA
                                                                                                               GAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAA
                                                                                                                          AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAA------AGGAGQGGYGGLG
                                                                                                                                                              SOGTSGREGLEGOGAGAAAAAA
                                                                                                                                                                                                     LGSQGTSGRGCLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT-----SGP
                                                                                                                                                                                                                                                                                                                                                                                     GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                GGQGAGAAAA----------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                      QGAGAAAAAAAAAAGGAGGGYGGLGSQGT---SGRGGLGGQGAG-AAAAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                         312 GGQGAGAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQG
                                                                                                                                                                                                                                                                                                           AGAAAAAAAAAAAGGGGYGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGG
                   Indels 176;
 Length 651;
; Score 2135; DB 4;
; Pred. No. 7.2e-151;
15; Mismatches 39;
61.1%;
68.3%;
                   Conservative
         Best_Local Similarity
Matches 495; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGLGS 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGVGS 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
US-08-425-069-2
Query Match
Best Local
                                                                                                                                                    200
                                                                          80
                                                                                             21
                                                                                                                140
                                                                                                                                  108
                                                                                                                                                                                          251
                                                                                                                                                                                                             202
                                                                                                                                                                                                                                 290
                                                                                                                                                                                                                                                   261
                                                                                                                                                                                                                                                                      349
                                                                                                                                                                                                                                                                                                           391
                                                                                                                                                                                                                                                                                                                             371
                                                                                                                                                                                                                                                                                                                                                                 420
                                                                                                                                                                                                                                                                                                                                                                                    498
                                                                                                                                                                                                                                                                                                                                                                                                       476
                                                                                                                                                                                                                                                                                                                                                                                                                           551
                                                                                                                                                                                                                                                                                                                                                                                                                                             532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       637
                                                                                                                                                                                                                                                                                                                                                451
                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                       g
                                                                          à
                                                                                            g
                                                                                                               ò
                                                                                                                               g
                                                                                                                                                    ò
                                                                                                                                                                    g
                                                                                                                                                                                         à
                                                                                                                                                                                                           g
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
```

```
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Xu, Mind
APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF NUMBER OF SEQUENCES: 62
CORRESPONDENCES: 62
ADDRESSEE: Birch, Stewart, Kolasch & Birch
                                                                                                  80 GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          140 GAAAAAAAAAAAAAAAAAAAAAA 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            532 GGGGAGAAAAAAGGAGGGLGGQGAGQGAGAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                ---- AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 176;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 718;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 61.1%; Score 2135; DB 2;
Best Local Similarity 68.3%; Pred. No. 7.8e-151;
Matches 495; Conservative 15; Mismatches 39;
                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5989894th Washington Street
CITY: Falls Church
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,844B
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE/DOCKET NUMBER: 1447-105P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                      ; Sequence 2, Application US/08317844B
; Patent No. 5989894
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEX: 248345
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Virginia
: U.S.A.
                                                                                                                                             651 GGLGS 655
                                                                                                                                                                           641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                          RESULT 4
US-08-317-844B-2
                                                                                                                                                               ||:||
|GGVGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STATE: V. COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-317-844B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEFAX:
                  551
                                                                                                                                                                           637
                                                                                                          g
                                               셤
                                                                              ò
                                                                                                                                           ò
                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGYGP-GQQTSGRGGLGGQGAGAAAAAAA ----AGGAGQGGYGGLGSQGTSGRGGL 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      200 AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA------AGGAGQGGYGGLG 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   251 SQGTSGRGGLGGQGAGAAAAAA-----------AAAAGGAGQGGYGGLGS 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           290 QGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LGSQGTSGRGCLGGQGAGAAAAAAAAAGAGQGGYGGLGSQGT-----SGP 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 QGAG----AAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGGQGAGAGGAGAAAAAAGGAGQ 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 61.1%; Score 2135; DB 1; Length 718; Best Local Similarity 68.3%; Pred. No. 7.8e-151; Matches 495; Conservative 15; Mismatches 39; Indels 176;
                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
: 301 No. 5728810th Washington Street
Falls Church
                                                                                                                                                                                                                ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGIGSTRATION NUMBER: 28,977
REFRENCE/DOCKET NUMBER: 1447-106P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
                                                                                                                                                                      APPLICATION NUMBER: US/08/425,069
FILLING DATE: 19-APR-1995
CLASSIFICATION: 435
                                                                                                                           PC-DOS/MS-DOS
                                                                                            MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                718 amino acids
                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear:
MOLECULE TYPE: protein
US-08-425-069-2
                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
                             Virginia
: U.S.A.
                                                           22046
                             STATE: V:
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

Gaps

Gaps

```
80 GCYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 139
                                                                                                                                                                                                                                                                                                                 312 GGGGAGAAAAGGAGGGGGGGGGGGAGAAAAAAGGAGGGGYGGLGSQG-AGRGGLGGQG 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23 QCAGAAAAAAAAGGAGGGGGGGGGGT--SGRGGLGGQGAG-AAAAAAAAAAGGAGO
                                                                                                                                                                                                                                                                                                                                                                                                 200 AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGQGAGAAAA----------AAAAAAGGAGGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       591 QGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                Indels 176;
                                                                                                                                                                                                          Length 747;
                                                                                                                                                                                                       Query Match 61.1%; Score 2135; DB 3; Best Local Similarity 68.3%; Pred. No. 8.1e-151, Matches 495; Conservative 15; Mismatches 39;
              as
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0486
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
                                                                   INFORMATION FOR SEQ ID NO: 3: SEQUENCE CHARACTERISTICS: LENGTH: 747 amino acids TYPE: amino acid STRANDEDNESS: single TOPOLOGY: linear INMEDIATE SOURCE: LIBRARY: Genbank
                                                                                                                                                                    GI 1174414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGVGS 641
                                                                                                                                                                    ; CLONE: C
US-09-034-177-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           637
                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                  201
                                                          289
                                                                                 260
                                                                                                           348
                                                                                                                                                        390
                                                                                                                                                                    391 AGAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGG 450
                                                                                                                                                                                                                 650
                                                                                                                    GGYGP-GQQTSGRGGLGGQGAGAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL 550
                                                                                                                                                                                                                                                                                                                                                  ---- AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 590
                                                                                                                                                                                                                                                                                                                                                                         587
                                                                                                                                                                                                                                                                                                                                                                                                                        636
                                                                                                                                                                                                                                                  LGSQCTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGT------SGP 497
                                                                                                                                                                                                                                                                                                                          531
        290 QGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                    349 GGQGAGAAAA------AAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                               ------AAAAGGAGGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                             591 QGAGAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                       532 GGQGAGAAAAAAGGAGGGLGGQGAGQGAGAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 3, Application US/09034177
Patent No. 6127146
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TTTLE OF INVENTION: HUMAN FIBROUS PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
COMPATING SYSTEM: DOS
COPFRATING SYSTEM: DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/034,177
                                                        251 SQGTSGRGGLGGQGAGAAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Billings, Lucy J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE: HEREWITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                GGQGAGAAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     641
                                                                                                                                                                                                                                                                                                                                                                                                                                             651 GGLGS 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZIP: 94304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGVGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 5
US-09-034-177-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE:
                                                                                                                                                                                                                                                  451
                                                                                                                                                                                                                                                                                                                                                551
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                g
                                                                             g
                                                                                                       ö
                                                                                                                          g
                                                                                                                                                  ò
                                                                                                                                                                        g
                                                                                                                                                                                                  à
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
         ò
                                                                                                                                                                                                                                                                                                 õ
```

250

289

348 311 390 450

497

550 531 587

```
AAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTS 419
                Ouery Match 59.4%; Score 2076; DB 4; Best Local Similarity 69.7%; Pred. No. 1.5e-146; Matches 477; Conservative 11; Mismatches 74;
                                                                                                                                                                                                                                                                AGGAGQGGYGGLGSQGTSGPGGYG 665
                                                                                                                                                                                                                                                                                                                           ; Sequence 4, Application US/09247806
; Patent No. 6280747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 606
                                                                                                                                                                                                                                                                                                                                                                                                     US-09-247-806-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local 8
                                                                                                                                                      117
                                                                                                                                                                                                                                                                                                           642
360
                                                           120
                                                                                                                                                                                                                 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       221
                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                      ద
                                                                                                                        ð
                                                                                                                                                   a
                                                                                                                                                                                 ò
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            221 AGQGGYGGLGSQG-AGRGGLGGQGAG----GGLGS 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       186 GQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT--SGRGGLGGQGA---GAAAAAAAAAAGG 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              241 AGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT--SGRGGLGG 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  137 QGA----GAAAAAAAAAAAGGAGQGGYGGLGSQGT--SGPGGYGPGQQT----SGRGGLG 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19 GLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAGGAG 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             74; Indels 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 606;
                                                                                                                                                           3: E. I. DU PONT DE NEMOURS AND COMPANY
1007 MARKET STREET
                                                                      GENERAL INFORMATION:

APPLICANT: PANNESTOCK, STEPHEN F.

TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED

TITLE OF INVENTION: SPIDER SILK ANALOGS

TITLE OF SEQUENCES: 107

CORRESPONDENCES: 107

CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               59.4%; Score 2076; DB 4;
69.7%; Pred. No. 1.5e-146;
ive 11; Mismatches 74;
                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CR-9389-A
                                                                                                                                                                                                              STATE: DELAWARE COUNTRY: UNITED STATES OF AMERICA
                                           Sequence 21, Application US/08556978B
Patent No. 6268169
                                                                                                                                                                                                                                                                                                                                                                                                                  08/077,600
                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,601
FILING DATE: JUNE 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION UNUMBER: 33,692
REFRENCE/CDOCKET NUMBER: CR-9;
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO: 21: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 606 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 477; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SS: unknown
unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       302-773-0164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                WILMINTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRANDEDNESS:
                                                                                                                                                                                                                                        19898
                                                                                                                                                                                                                                                                                                                                                                     FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-556-978B-21
                                                                                                                                                                                 STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
```

```
30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: PHILLIPPE, Michel
APPLICANT: GRASON, Jean-Claude
APPLICANT: GRASON, Jean-Claude
APPLICANT: ARRADDEAU, Jean-Platre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER PELING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 14
                        --GGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGGGGGGGGGGGGGGGGGGGGGAAAAAA416
                                                                                                                                    137 QGA---GAAAAAAAAAAGGAGGGYGGLGSQGT--SGPGGYGPGQQT-----SGRGGLG 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               527 AAGGAGQGGYGGLGSQGT--SGRGGLGGQGA---GAAAAAAAAAAGAGGGGYGGLGSQG 581
                                                                                                                                                                                                                                                                                                               582 TSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 641
                                                                                                                                                                                                                                                                                                                                                             -----GLGSQGAGQGAGAAAA 578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          79 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT--SGRGGLGG 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               241 AGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----GGLGS 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19 GLGGQGAGAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAG 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAAGGAGGGYGGLGSQGT--SGPGGYGPGQQT----SGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                         GRGCLGCQGAGAAAAAAAAAAAGGGGGGGGGGGGGGGGGT--SGRGCLGCQGA---GAAAAAA
```

9

us-09-490-291-4.rai

atch 56.5%; Sco cal_Similarity 64.2%; Pre
Malches 402; Conservative B; M
_
UD 4.3 SQUAGQAGAAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGG
Qy 196 AAAAAGGAGGGKGGSQGTSGRGGLG
Db 154 AAAAAGGAGQGGYGGLGSQGAGRG
206
0.y 304 GAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGG
364
Qy 412 GLGSQGTSGRGGLGGQGAGAAAAAAA
Db 342GGLGSQGAGARAAAA
Qy 472 AAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
Qy 523AAAAAAGGAGQGCYG
Db 443 GGLGSQGAGQGAGAAAAAAGGAGQGGYG
Qy 570 GOGGYGGLGSOGTSGR
:
Qy 599AAAAAAGGAGQGGYGGLGSQGTSGRG
Db 555 AGAAAAAGGAGQGCYGGLGSQGA
0 1111000
US-09-247-806-8
; Patent No. 6280747
; GENERAL INFORMATION: ; APPLICANT: PHILLIPPE, Michel
; APPLICANT: GARSON, Jean-Claude ; APPLICANT: ARRAUDEAU, Jean-Pierre
; TITLE OF INVENTION: COSMETIC OR DER
; TITLE OF INVENTION: ANALOG
CURRENT APPLICATION NUMBER: US/09/2 CURRENT FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: F 98/0 . PARTIER FILING DATE: 1008-00-11
NUMBER OF SEQ ID NOS: 14 SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8 ; LENGTH: 606

```
29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ERMATOLOGICAL COMPOSITION CONTACTING AT
JRAL OR RECOMBINANT SPIDER SILK OR AN
                                                                                                                                                                                   AAAAAAAGCAGQGGYGGLGSQGTSGRGGLGGQGA 303
                                                                                                                                                                                                                                                                             ISGPGGYGPGQQTSGRGGLGGGGAGAAAAAAA 363
                                                                                                                                                                                                                                                                                                                           CGCLGSQGTSGRGGLGGQGAGAAAAAAAAAGGA 569
||||||||||
|CGLGSQG----AGRGGQGAG----AAAAAAGGA 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3LGGQGAGAAA-----A 195
                                                                                                                                                       AAAAAAGGAGQ------GGLGSQGA 248
                                                                                                                                                                                                                                                                                                                                                                                                                     GGYGP-GQQTSGRGGLGGQGAGAAA----- 522
                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAA 471
                                                                                                                                                                                                                                                                                                                                                                                     3RGGLGGQGAGAAA------ 598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BRGGLGGQGAGAAAAAAGGAGQGGLGSQGAGQG 554
                      Indels 215; Gaps
                                                          SAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 97
Score 1972.5; DB 4; Length 606;
Pred. No. 6.7e-139;
Hismatches 35; Indels 215;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             747,806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    701614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGIH: 606
```

ŏ

ð

```
208 QGGYGGLGSQGAGRGGLGGQGAG---AAAAAAAGGAGG------GGLGSQGA 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA---------A 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----SGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      364 AG-----GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQ 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             389 AAGGAGQGGYGGLGSQG-AGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAAGGGGLGS 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     196 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGGGYGGLGSQGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels 207;
                                                                                                                                                                                                                                                                                                                                                                                                                  Length 604;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  477 AAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGLGSQGT-----SGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                  DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                               Score 1971.5; DB 4.
Pred. No. 7.9e-139;
8; Mismatches 37;
                      COMPOTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS SOFTWARE: MICROSOFT WINDOWS SOFTWARE: MICROSOFT WINDOWS CURRENT APPLICATION DAYR:
APPLICATION NUMBER: US/08/556,978B FILING DATE: 08/077,600
FILING DATE: UNMER: 08/077,600
FILING DATE: UNMER: 1993
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD. LINDA AXAMETHY
NAME: FLOYD. LINDA AXAMETHY
REFERENCE/DOCKET NUMBER: 33.692
REFERENCE/DOCKET NUMBER: CR-9389-A
                                      MICROSOFT WINDOWS 9
               DISKETTE, 3.50 INCH
                                                                                                                                                                                                                                TELECOMMUNICATION INFORMATION
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
                                                                                                                                                                                                                                                                                                                                                                                                                56.4%;
                                                                                                                                                                                                                                                                                                          604 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                TELEFAX: 302-773-0164 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: unknown
MOLECULE TYPE: protein
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                         TYPE: amino acid
STRANDEDNESS: un)
                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                   US-08-556-978B-63
                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 461;
                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                     29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        289 AGGAGQGGYGGLGSQGAGGGGGGGGGGGGGGGGG---- 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGA 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA---------A 195
                                                                                                                                                                                                                                                                                                                      -----SGRGGLGGQGAGAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGA 303
                                                                                                                                                                                                                                                                                                                                                                                  206 QGGYGGLGSQGAGRGGLGGQGAG---AAAAAAAGGAGQ--------GGLGSQGA 248
                                                                                                                                                                                                                                                                                                                                                                                                                GOGAGAAAAAAGGAGOGGYGGLGSQG------AGRGGOGAG---AAAAA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAAGGGGGYG 411
                                                                                     Gaps
                                                                                                                                94
                                                                                                               GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   412 GLGSQGTSGRGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         472 AAAAAAAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGAAA-----
                                                                                                                                                                                                      43 SQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                            GOGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAAAAAGGGGLGSQGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --AAAAAAAGAGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQ
                                                         Length 606;
                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            570 GQGGYGGLGSQGT-----SGRGGLGGQGAGAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 63, Application US/08556978B
Patent No. 6268169
GENERAL INFORMATION:
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
TITLE OF INVENTION: SPIDER SILK ANALOGS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
                                                         DB 4;
                                                      Score 1972.5; DB 4
Pred. No. 6.7e-139;
8; Mismatches 35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: WILMINTON
STATE: DELAWARE
COUNTX: UNITED STATES OF AMERICA
ZIP: 19898
                                                      56.5%;
; TYPE: PRT
; ORGANISM: Nephila clavipes
US-09-247-806-8
                                                                                     Conservative
                                                                    Best Local Similarity
Matches 462; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-556-978B-63
                                                         Query Match
                                                                                                                38
                                                                                                                                                                                                                                                                                                                                                                                                               304
                                                                                                                                                                                                                                                                                                                                                                                                                                            249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         495
                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
```

ò

ò

.

 SOO GGLGSQGAGGGGYGGLGSQGAGGGGGGAGAAAAAAAAAAAAAAAAAAAAAAAA	
Oy 602 AAAGGACGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGGGGYGGLG 654	OY 570 GQGGYGGLGSQGTSGRGCLGGQGAGAAA 1
RESULT 11 US-09-247-806-6 S Sequence 6, Application US/09247806 Patent NO. 6280747	Qy 599AAAAAAGGAGGGYGGLGSQGTSGRGGLGGGGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
GENERAL INCORPORTION: APPLICANT: PHILLIPPE, Michel APPLICANT: GARSON, Jean-Claude APPLICANT: ARRAUDEAU, Jean-Pierre TITLE OF TWVENTYON: COGMETT OF DEPART	RESULT 12 US-08-209-747-2 ; Sequence 2, Application US/08209747 ; Patent No. 5733771
TITLE OF INVENTION: CARACTURE OR DEMANDED COMPOSITION CONTROLLING AT TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN TITLE OF INVENTION: ANALOG : FILE REFERENCE: 6388-0365-0 : CURRENT APPLICATION NUMBER: US/09/247,806 : CURRENT FILING DAME: 1999-02-11	GENERAL INFORMATION: APPLICANT: Lewis, Randolph V. APPLICANT: Colgin, Mark TITLE OF INVENTION: CDNAS Encoding Minor Ampulls TITLE OF INVENTION: SILK Proteins
) NUMBER OF SEQUENCES: 56 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: Birch, Stewart, Kolasch & Birch ; STREET: P.O. BOx 747 ; CITY: Falls Church
; SEU ID 0 5 ; LENGTH: 606 ; TYPE: PRT ; ORGANISM: Nephila clavipes US-09-247-806-6	; STATE: Virginia ; COUNTRY: USA ; ZIP: 22040-3487 ; COMPUTER READABLE FORM: ; MEDIUM TYPE: Floody disk
Query Match 56.2%; Score 1961.5; DB 4; Length 606; Best Local Similarity 64.1%; Pred. No. 4.3e-138; Matches 460; Conservative 8; Mismatches 35; Indels 215; Gaps 29;	COMPUTER: IBM PC Compatible COMPRATION SYSTEM: PC-DOS/MS-DOS COFFMARE: PatentIn Release #1.0, Version #1.2. CURRENT APPLICATION DATA:
QY 38 GAGGGGYGGLGGSQGTSGRGGLGGGGGGAGAAAAAAAAAAAGGAGGGGGGGGGGGG	; APPLICATION NUMBER: US/08/209,747 ; FILING DATE: 14-MAR-1994 ; CLASSIFICATION: 530 ; ATVONEY/AGENT INFORMATION:
Qy 98 GQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGAGGGG 157 Db 45 SQGAGQGAGAAAAAAGGAGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGG 96	; NAME: Murphy Jr., Gerald M. ; REGISSTRATION NUMBER: 28,977 ; REFERENCE/DOCKET NUMBER: 1447-104P ; TELECOMMUNICATION INFORMATION:
Qy 158 YGGLGSQGTSGPGGYPGGCGGQGAGAAAA	; TELEPHONE: 703-205-8000 ; TELEPAX: 703-205-8050 ; INFORMATION FOR SEQ ID NO: 2: ; SEQUENCE CHARACTERISCS:
Qy 196 AAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGGGYGGLGSQGT- 254 D 156 AAAAAGGAGGGYGGLGSQGAGRGGQGAGAAAAAAGGAGGGYGGLGSQGAG 207	; LENGTH: 832 amino acids ; TYPE: amino acid ; STRANDEDNESS: single ; TOPOLOGY: linear
Oy 255SGRGGLGGGGAGAAAAAAAAGGGGGGGGGGGGGGGGGGGG	E I E
304 GAAAAAAAAGGAGGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA) OKRANISM: N. Clavipes ; TISSUE TYPE: minor ampullate gland ; FEATURE:
DD 251 GQGAGAAAAAAGGAGQGGGGGGGGGGGGGGGGGGGGGG	; NAME/KEY: CDS ; LOCATION: 1309 US-08-209-747-2
Db 291 AGGAGGGGYGGLGSQGACQGGYGGLGSQGAGRGGLGGQGAAAAAAAAAAGGGGQ 343 Qy 412 GLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGGGGGGGGGGGGGGGGGG	Query Match 40.2%; Score 1405; DB 1; Best Local Similarity 50.1%; Pred. No. 6.7e-97; Matches 349; Conservative 17; Mismatches 207;
Db 344GGLGSQGAGAAAAAAGGGGGGGGGGGGGGGGAGRGGQGAG 386	Oy 21 GGGGAGAAAAAAAGGAGGGGYGGLGSGGTGGGGGGGGG
AAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	79
523	Qy 81 GYGGLGSQGTSGRGGLGGQGAQAAAAAAAGGGGGGGG

```
26;
865 ----- 288
                            AGGAGQGGLGSQGAGQG 556
                                                                                                                                                                                                                                                                                                                                                                                              Indels 124; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                 ---GLGSQGTSGRGGLG 135
| : :| || |
NGAGAAAAGAGAGGAG 183
                                                                                                                                                                                                                                                                                                                                                                                                          Length 832;
                                                                                                                 ate Spider
```

```
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                    ----LGSQGTSGRGGLG---GQGAGAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPG 339
                                                                                                                                                                                                                                                      470
                                                                                                                                                                                                                                                                                    448
GQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAA 195
                                                                                                                               354 GAAAAAGAGAGGAGGYGRGAGAGAGAAAAAGAGAAAAAGAGAGGYGGQGGYG-AGAGAGAAA 412
                                                                                                                                                                                                                                                                                                                                          503
                                                                                                                                                                                                                                                                                                                                                                                                  554
                                                                                                                                                                                                                                                                                                                                                                                                                              637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        521
                                                                                                                                                                                                                                                                                                                                                                   522 GRGAG----AGAGAAAGAGAGAAGAGAGGYGGYGGYGAGARAGAAAAAGAGAGGAAG 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                       AGAAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQCAGAAAAAAAAAAAAGGAGQGGYGG 614
                                                                                                                                                                                                                                            QGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                                                                 449 G-GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGG----LGSQGTSGPG-GYGPG
                                                                                                                                                                                                                                                                                                                                                                                                504 QQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYG---GLGS-----QGTSGRGGLG-GQG
                                                                                  242 GAAAGAGAGAGAGSYG------GQGYGAGAGAAAAAXXXXXXXXXXXXXXXXXXAGAGAG
                                                                                                            240 GAGQGGYG-----GLGSQGTSGRGCLGGQGAGAAAAAAAAAGGAGQGGYGG-----
                                                                                                                                                                                                                          340 QQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQG-----TSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
APPLICANT: Colgin, Mark
TITLE CANTENTION: CDNAS Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. BOX 747
CITY: Falls Church
STATE: Virginia
COUNTER: USA
ZIP: 22040-3487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             615 LGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYG 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             196 AAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,747
APLING DATE: 14-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/08/458,298 FILLING DATE: 02-JUN-1995 CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 2, Application US/08458298
Patent No. 5756677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-458-298-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                       555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   638
136
                                                                                                                                                                                                                                                                                  389
                                                                                                                                                                    287
                       g
                                                      ò
                                                                               셤
                                                                                                                                  a
                                                                                                                                                                                             g
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                            ò
                                                                                                                                                                 ò
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
```

```
GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYG----GLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           136 GQCAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAA 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            242 GAAAGAGAGAGSYG------GQGYGAGAGAAAAAXXXXXXXXXXXXXXAGAGAG 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        196 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA-------AAAAAG 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21 GGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQG 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       449 G-GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG---LGSQGTSGPG-GYGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        240 GAGQGGYG----GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   504 QQTSGRGGLGGQGAGAAAAAAAAAAAAGGAGGGGYG---GLGS-----QGTSGRGGLG-GQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    578 YSRGGRAGAAGAGAAAGAGAGAGGYGGGGGGGGAGAGAAAAAGAGSGGAGGYGRGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                340 QQTSGRGGLGGQGAGAAAAAAAAAAAGAGGAGGCYGGLGSQG------TSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 40.2%; Score 1405; DB 1; Length 832; Best Local Similarity 50.1%; Pred. No. 6.7e-97; Matches 349; Conservative 17; Mismatches 207; Indels 124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            515 LGSQGTSGRGCLGCQCAGAAAAAAAAAGGAGQGGYG 551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
TISSUE TYPE: minor ampullate gland
                  REFERENCE/DOCKET NUMBER: 14
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEPAX: 703-205-8050
                                                                                 TELEFAX: 703-205-8050 INFORMATION FOR SEQ ID NO: 2:
                                                                                                                     SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
                                                                                                                                                                               STRANDEDNESS: single
TOPOLOGY: linear
REGISTRATION NUMBER:
                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                      ; LOCATION:
US-08-458-298-2
                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY:
```

556 593 611

999

```
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Tlaude
APPLICANT: ARRADDRAU, Jean-Plerre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ARALOG
FILE REFERENCE: 6388-0355-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
                          360 SGPGSAAAAAA-GPGQQGPGGYG-PGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAGPGGYGP 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66 AAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAGGGGGY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 GGLGSQGTSGRGGLG-GQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180 GRGGLGGQGAGAAAAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAA 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21 GGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAGPGGYGPGQQGPGGYGPGQQGPSGPGS 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 GSMASGRGGLGGQGAG----AAAAAAAAAAAGGAGGGGYG-GLGSQGTSGRGGLGGQGAGA 65
                                                                                                                              -GOOGPGGYGP-GOOGPSGPGSAAAAAAAAAGPGGYGP-GOOGPGGYGPG
                                                                                                                                                                                                                                                                                     475 QQGPSGPGSAAAAAAQPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAGPGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                 YG-GLGSQGTSGRGGLGGQGAGAAAAAAAAAG----GAGQGGYGGLGSQGTSGPGGYGP
354 GAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAAAGGAGQGGYGG
                                                                                      LGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG----GLGGQGAG
                                                                                                                                                                        ------AAAAAAAAAAGGGGGYG------GLGSQGTS------GPGGY
                                                                                                                                                                                                                                                        501 GPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG----GLGGQGAG
                                                                                                                                                                                                                                                                                                                                               Query Match 38.1%; Score 1331.5; DB 4; Length 714; Best Local Similarity 50.2%; Pred. No. 1.6e-91; Matches 363; Conservative 25; Mismatches 242; Indels 93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OTHER INFORMATION: Description of Artificial Sequence:protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 10, Application US/09247806 Patent No. 6280747 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     699 005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             712 G00 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-247-806-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-247-806-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURE:
                                                                                                                              418
                                                                                                                                                                                                                                                                                                                                               557
                                                                                                                                                                                                                                                                                                                                                                                                                                 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     299
                                                                                                                                                                        469
                                          염
                                                                                    ò
                                                                                                                          g
                                                                                                                                                                        ò
                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81 AAAAAAAAAGPGGYGPGQQGPGGYG-PGQQGPGGYGPGQQGPSGPGSAAAAAAGPGQQGP 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 GGLGSQGTSGRGGLG-GQGAGAAAAAAAAGGAGGGGYGGLGSQGTSGPGGYGPGQQTS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        140 GGYG-PGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAAAGPGGYG----PGQQGPGGYGPGQQ-- 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----GPSGPGSAAAAAAAAAAGPGGYGPGQYGPGGYG-PGQQGPGGYGPGQQGPSGPGSA 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 294 GRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGA 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GYGP-GQQGPSGPGSAAAAAAAAAAAAGPGGYG----PGQQGPGGYGPGQQGPGGYGPGQQGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11 GSMASGRGGLGGQGAG----AAAAAAAAAAAGGAGGGYG-GLGSQGTSGRGGLGGQGAGA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21 GGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAGPGGYGPGQQGPGGYGPGQQGPSGPGS 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     235 AAAAGGAGGGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAAGGAGGGGYGGLGSQGTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                180 GRGGLGGQGAGAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 714;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38.1%; Score 1331.5; DB 4; Length 50.2%; Pred. No. 1.6e-91; Live 25; Mismatches 242; Indels
                                                                                                                                                                                                                                  :: E. I. DU PONT DE NEMOURS AND COMPANY
1007 MARKET STREET
                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
                                                                                                                                         TITLE OF INVENTION: NOVEL RECOMBINANTLY TITLE OF INVENTION: SPIDER SILK ANALOGS NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9389-A
TELECOMMUNICATION INFORMATION:
TELEPAX: 302-892-8112
TELEFAX: 302-773-0164
FORMATION FOR CRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
                                                                                                                                                                                                                                                                                                                    UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,600
FILING DATE: JUNE 15, 1993
ATTORNEY/AGBNT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
                                                           Sequence 61, Application US/08556978B Patent No. 6268169
                                                                               Patent No. 6268169
GENERAL INFORMATION:
APPLICANT: FAHNESTOCK, STEPHEN F.
                                                                                                                                                                                                                                                                                                                                                                                                      IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : 714 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 50.29
Matches 363; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: unknown
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                STATE: DELAWARE
                                                                                                                                                                                                                                                                          WILMINTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                  ADDRESSEE:
STREET: 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-556-978B-61
                                   -08-556-978B-61
                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

35;

Gaps

```
294 GRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGA 353
                                      354 GAAAAAAAAAAAGGAGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAAGGAGQGGYGG 412
                                                         360 SGPGSAAAAA-GPGQQGPGGYG-PGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYGP 417
                                                                      534
                                                                                                            501 GPGQQTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG----GLGGQGAG 556
                                                                                                                  699 000 299
                                                                                                                                                                             G00 714
                                                                                          469
                                          305
                                                                                                                            P G
              ò
                     8 6 6 6
6 6 6
                                                          ò
                                                                                                                                                         qq
                                                                                                                                                                    ò
```

Search completed: December 18, 2002, 16:23:06 Job time: 18.6792 secs

This Page Blank (uspto)

```
December 18, 2002, 16:21:10 ; Search time 9.40752 Seconds (without alignments) 1213.493 Million cell updates/sec
                                                                                                                                                                                                                                                                                      US-09-490-291-4
3493
1 MRGSHHHHHHGSMASGRGGL......GGYGPGQQTSGIRRPAAKLN 681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/USOG_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/USOC_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/USOC_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/USOC_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/USOC_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 106657
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    106657 seqs, 16763532 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                 OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0 Maximum DB seq length: 20000000000
                                                                                                                                                                                                                                                                                             Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Searched:
                                                                                                                                                                                                                                                                                                                                                       Sequence:
                                                                                                                                                                        Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	Sequence 1, Appli	Sequence 4, Appli	Sequence 8, Appli	Sequence 6, Appli	Sequence 10, Appl	Sequence 2, Appli	Sequence 97, Appl	Sequence 2, Appli	Sequence 269, App	Sequence 2, Appli		Sequence 1, Appli	Sequence 37, Appl	Sequence 20, Appl	Sequence 525, App	Sequence 35807, A	Sequence 451, App	Sequence 1, Appli	3,
SUMMARIES	ΠD	US-09-861-597-1	US-09-861-597-4	US-09-861-597-8	US-09-861-597-6	US-09-861-597-10	US-09-861-597-2	US-10-052-586-97	US-09-804-898-2	US-10-052-586-269	US-09-060-854B-2	US-10-029-217A-24	US-10-096-961-1	US-09-879-957-37	US-09-756-071B-20	US-10-052-586-525	US-09-864-761-35807	US-10-052-586-451	US-09-735-933-1	US-09-826-508-3
		10	10	10	10	10	10	12	10	12	10	<u>о</u>	6	10	10	12	10	12	10	10
	% Query Watch Length DB	651	909	909	909	714	529	1894	4679	1300	1497	3907	2211	1400	720	4440	357	1743	1422	1200
	& Query Match	61.1	59.4	56.5	56.2	38.1	31.4	24.7	21.1	19.6	19.3	19.2	18.8	18.7	18.7	17.4	17.4	17.3	16.7	16.3
	Score	2135	2076	1972.5	1961.5	1331.5	1097.5	861.5	738	989	675.5	672	656.5	652	651.5	609	607	605.5	582	570.5
	Result No.	1	7	3	4	S	9	7	œ	σ	10	11	12	13	14	15	16	17	18	19

45 A

Sequence 47, Appl Sequence 581, Appl Sequence 581, Appl Sequence 36720, A An-37 Sequence 37, Appl Sequence 37, Appl Sequence 37, Appl Sequence 34, Appl Sequence 34, Appl Sequence 34, Appl Sequence 34, Appl Sequence 39, Appl Sequence 35, Appl Sequence 36, Appl Sequence 14, Appl Sequence 14, Appl Sequence 18, Appl Sequence 25, Appl Sequ	RESULT 1 US-09-861-597-1 US-09-861-597-1 Sequence 1, Application US/09861597 SEQUENCE 1, UNIVERSITY OF TAILLIPPE, Michel APPLICANT: PHILLIPPE, Michel APPLICANT: GARSON, Jean-Pierre APPLICANT: USAGNETIC OR DERMATOLOGICAL COMPOSITION CONTACTING TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN TITLE OF INVENTION: ANALOG CURRENT FILING DATE: 1999-02-11 PRIOR APPLICATION NUMBER: CR 98/01614 PRIOR FILING DATE: 1999-02-11 PRIOR FILING DATE: 1999-02-11 PRIOR FILING DATE: 1999-02-11 PRIOR FILING DATE: 1999-02-11 SEQ ID NO 1 LENGTH: 651 TYPE: PRT CORGANISM: Nephila clavipes US-09-861-597-1	tch 61.1%; Score 2135; DB 10; Length 651; 495; Conservative 15; Mismatches 39; Indels 176; Gaps 34; QGAGAAAAAAAAGGAGGGGGGGGGGGGGGGGGGGGGA-AAAAAA
US-09-975-143-47 US-10-052-586-581 US-09-919-497-56 US-09-819-497-56 US-09-817-9659-34 US-09-817-9659-34 US-09-817-9659-34 US-09-817-9659-34 US-09-817-9659-34 US-09-817-9659-34 US-09-817-9659-36 US-09-817-9659-36 US-09-817-9659-36 US-09-817-9659-36 US-09-817-9659-36 US-09-8186-055-431 US-09-866-055-431 US-09-866-055-431 US-09-817-9659-18 US-09-817-9659-18 US-09-817-9659-18 US-09-817-9659-18 US-09-817-9659-18 US-09-817-9659-18 US-09-817-9659-18 US-09-817-9659-18	ALIGNMENTS 861597 ude Plerre OR DERMATOLOGIC NE NATURAL OR RE US/09/861,597 5-22 747,806 1	; Score 2135; Pred. No. 1e- 15; Mismatches YGGLGSQGTSGRGG
7.7.7.7.7.7.7.7.7.7.7.7.7.7.7.7.7.7.7.	Cation US/098615; 006453941 ON: IPPE, Michel ON, Jean-Pierr ON: COSMETIC OR I ON: LEAST ONE NO. ON: LOON UNBER: US/01 ANDER: 2001-05-22 N NUMBER: 09/247 E: 1998-02-11 IN VET. 2.1 In Vet. 2.1	larity
25	SULT 1 Sequence 1, Application Sequence 1, Application Sequence 1, Application Septican No. USCO06453 APPLICANT: PHILLIPPE, APPLICANT: ARRADEBAU, TITLE OF INVENTION: COTTILE OF INVENTION: COTTILE OF INVENTION: COTTILE OF INVENTION: LITTLE OF INVENTION: COURRENT APPLICATION NUMB PRIOR FILING DATE: 199 PRIOR FILING DATE: 199 PRIOR FILING DATE: 199 RIOR FILING RIOR	Query Match Best Local Similarity Matches 495; Conser 23 QGAGAAAAAAAA 1 1 QGAGAAAAA 7 80 GGYGGLGSQGTSG 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
0.10202020202020202020202020202020202020	RESULT 1 US-09-86; Sequency Partial Result R	Query Best Matcl Qy Db Qy

```
US-09-861-597-8
  m
                  2
                                  55
                                                                                   186
                                                                                                                                                   301
                                                                                                                                                                                  360
                                                                                                                                                                                                                                                   475
                                                                                                                                                                                                                                                                  417
                                                                                                                                                                                                                                                                                                   477
                                                                                                                                                                                                                                                                                   527
                  ò
                                 셤
                                                 ò
                                                               g
                                                                                  à
                                                                                                 g
                                                                                                                δ
                                                                                                                                g
                                                                                                                                                  à
                                                                                                                                                               g
                                                                                                                                                                                  ŏ
                                                                                                                                                                                                 g
                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                    Sequence 4. Application US/09861597

Fatent No. US20020064539A1

GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: ARRAUDBAU, Jean-Pleire
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT PILING DATE: 1999-02-11
PRIOR APPLICATION NUMBER: FR 98/01614
PRIOR FILING DATE: 1998-02-11
PRIOR FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SSOTUMENT PATCHIN VET. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30;
                                                                                                                                  450
 250
                  201
                                        QGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL 348
                                                                                                          GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL 550
                                                                                                                                                                                                                   531
                                                                                                                                                                                                                                                  587
                                                                                                                                                                                                                                                                  650
                                                                                                                                                                                                                                                                                   636
                                                                         Gaps
-----AAAAGGAGQGGYGGLGS
                                                                                                 GCQGAGAAAA------AAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                 Indels 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Score 2076; DB 10;
; Pred. No. 2.2e-112;
11; Mismatches 74;
                                  SQGTSGRGGLGGQGAGAAAAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 59.4%;
Best Local Similarity 69.7%;
Matches 477; Conservative 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; ORGANISM: Nephila clavipes US-09-861-597-4
                                                                                                                                                                                                                                                                                                           ||:||
GGVGS 641
                                                                                                                                                                                                                                                                                                 GGLGS 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 606
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                  US-09-861-597-4
                                                 202
                                                                                 261
                                                                                                 349
                                                                                                                                 391
                                                                                                                                                                                                                 476
                                                                                                                                                                                                                                  551
                                                                                                                                                                                                                                                 532
                                                                                                                                                                                                                                                                                                 651
                                                                                                                                                                                                                                                                                                                  637
                                                                 290
                                                                                                                                                 371
                                                                                                                                                                 451
                                                                                                                                                                                 420
                                                                                                                                                                                                  498
               a
                                               g
                                                                à
                                                                               g
                                                                                                ö
                                                                                                              g
                                                                                                                                 ò
                                                                                                                                               g
                                                                                                                                                              ò
                                                                                                                                                                                g
                                                                                                                                                                                                 ö
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            å
ð
                                 à
```

```
Sequence 8 Application US/09861597

Sequence 8 US20020064539A1

SEQUENCE INFORMATION:
SEQUENCE INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: ARRAUBEAU, Jean-Pierre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT APPLICATION NUMBER: 09/247,806
PRIOR APPLICATION NUMBER: PR 98/01614
PRIOR APPLICATION NUMBER: FR 98/01614
PRIOR FILING DATE: 1998-02-11
SEQ ID NO 8

LENGTH: 606

LENGTH: 606

LENGTH: 606
                                                                                                            185
                                                                                                                                              160
                                                                                                                                                                                 240
                                                                                                                                                                                                                     220
                                                                                                                                                                                                                                                    241 AGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGG 300
                                                                                                                                                                                                                                                                                                                           314
                                                                                                                                                                                                                                                                                                                                                                                                 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GRGGQGAG----AAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAG---AAAAAAAGGAG
                                                      GQGAGAAAAAAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGA---GAAAAAAAAAGG
                                                                                                                                                                                                  AAAAGGAGGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGGGGLGSQGAGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT - - SGRGGLGG
                                                                                                          137 QCA---GAAAAAAAAAAAGAGGGYGGLGSQGT--SGPGGYGPGQQT----SGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                   AAAAGGAGQGGYGGLGSQGT--SGPGGYGPGQQT----SGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAGGAGQGGYGGLGSQGT--SGRGGLGGQGA---GAAAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         582 TSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            642 AGGAGQGGYGGLGSQGTSGPGGYG 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; TYPE: PRT
; ORGANISM: Nephila clavipes
US-09-861-597-8
```

Gaps

Indels

Length 606;

4

97

195

96

254 207 303 250 411

471

```
3 GAGQGGYGGLGSQG-AGRGGLGGQGAG---AAAAAAAGGAGG---------GGLG
                                                                                                                                                                                                                                                                                                               108 QGGYGGLGSQGAGRGGLGGQGAG-------GGLGSQGA
                                                                                                                                                                                                                                                                                                                                                                                                       570 GQGGYGGLGSQGT-----SGRGGLGGQGAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        38 GAGOGGYGGLGSQGTSGRGGLGGOGAGAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLG
                                                                                                                                                                                  45 SQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AAAAAAAGGAGQGG
                                                                                                                                                                                                                                                     196 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-
                                                                                                                                                                                                                                                                                                   364 AGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                       472 AAAAAAAGGAGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAGGA
                                                                                                                                                                                                         158 YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA------A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        599 -- AAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLG
                                                                        DB 10;
                                                                    56.2%; Score 1961.5; DB 1
llarity 64.1%; Pred. No. 7.3e-106;
Conservative 8; Mismatches 35;
                                   ; ORGANISM: Nephila clavipes US-09-861-597-6
                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 5
US-09-861-597-10
   NO 6
PH: 606
                                                                                 Best Local Sim
Matches 460;
                           TYPE: PRT
                                                                        Query Match
              LENGTH:
                                                                                                                                                                                                                                                                                                   255
                                                                                                                                                                                                                                                                                                                                               304
                                                                                                                                                                                                                                                                                                                                                                                                                                                            344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              557
   SEQ ID
                                                                                                                  à
                                                                                                                                     셤
                                                                                                                                                              à
                                                                                                                                                                                 ä
                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                  à
                       29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: ARRADBAU, Jean-Pietre
APPLICANT: ARRADBAU, Jean-Pietre
APPLICANT: ARRADBAU, Jean-Pietre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: ALLOG
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: 09/247,806
PRIOR FILING DATE: 1999-02-11
PRIOR APPLICATION NUMBER: FR 98/01614
                                                                                                                                       ------ 195
                                                                                                                                                  522
                                                                                                                                                                                                                                                                                                                                                                                                                           206 QGGYGGLGSQGAGRGGLGGQGAG---AAAAAAAAGGAGQ-------GGLGSQGA 248
                                                                                                                                                                                                                                                                                                                                   ------AAAAAAGGAGQGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAAGGA 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GQGGYGGLGSQGT------SGRGGLGGQGAGAAA-------598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     929
                                                                                                                                                                                  254
                                                                                                                                                                                                                                                                         GQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAAAAAGGAGQGG 157
                          Gaps
                                                           38 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT-
                                                                                                                                                                                                                                                                                                                                                                                                               472 AAAAAAAGGAGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-----
 Score 1972.5; DB 10; Length 606;
Pred. No. 1.7e-106;
8; Mismatches 35; Indels 215;
                                                                                                                                     158 YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 6, Application US/09861597
Patent No. US20020064539A1
GENERAL INFORMATION:
56.5%;
64.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
                       Matches 462; Conservative
             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     us-09-861-597-6
 Query Match
              Local
                                                                                          86
                                                                                                                                                            95
                                                                                                                                                                                 196
                                                                                                                                                                                                                              255
                                                                                                                                                                                                                                                                         304
                                                                                                                                                                                                                                                                                                                                             289
                                                                                                                                                                                                                                                                                                                                                                                        342
                                                                                                                                                                                                                                                                                                                                                                                                                                                           523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       570
                                                                                       à
                                                                                                             셤
                                                                                                                                    ò
                                                                                                                                                          g
                                                                                                                                                                                 ò
                                                                                                                                                                                                   Q
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
Sequence 10, Application US/09861597
Patent No. US200064539A1
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Pleare
APPLICANT: GRSON, Jean-Pleare
APPLICANT: GRSON, Jean-Pleare
APPLICANT: GRSON, Jean-Pleare
APPLICANT: ANALOG
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6198-01365-0
CURRENT APPLICATION UNBER: US/09/861,597
CURRENT FILING DATE: 2001-05-22
```

```
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81 AAAAAAAAAGPGGYGPGQQGPGGYG-PGQQGPGGYGPGQQGPSGPGSAAAAAAAGPGQQGP 139
                                                                                                                                                                                                                                                                                                                                                                                         121 GGLGSQGTSGRGGLG-GQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTS 179
                                                                                                                                                                                                                                                                                                                                                                                                                     140 GGYG-PGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYG----PGQQGPGGYGPGQQ-- 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                    180 GRGGLGGQGAGAAAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAA 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               235 AAAAGGAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAGGAGGGGGGGGGGGGTS 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGA 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  354 GAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAAGGAGQGGYGG 412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           413 LGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGGGYGGLGSQGTSGRG----GLGGQGAG 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   594 QQGPSGPGSAAAAAAGPGQQGPGGYG-PGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAGPGG 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               360 SGPGSAAAAA-GPGQQGPGGYG-PGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYGP 417
                                                                                                                                                                                                                                                                                                   21 GGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYGPGQQGPGGYGPGQQGPSGPGS 80
                                                                                                                                                                                                                                                                       11 GSMASGRGGLGGQGAG----AAAAAAAAAAAGGAGGGGGG-GLGSQCTSGRGGLGGQGAGA 65
                                                                                                                                                                                                                                                                                                                                66 AAAAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAAAGCAGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          475 QQGPSGPGSAAAAAAGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAGPGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              501 GPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG----GLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         557 ----AAAAAAAAAAAAAAAAAGGGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      469 ------AAAAAAAAAGGAGGGYG------GLGSQGTS-------GPGGY
                                                                                                                                                                                                              Query Match 38.1%; Score 1331.5; DB 10; Length 714; Best Local Similarity 50.2%; Pred. No. 6.3e-70; Matches 363; Conservative 25; Mismatches 242; Indels 93;
                                                                                                                                                                  ; OTHER INFORMATION: Description of Artificial Sequence:protein US-09-861-597-10
PRIOR APPLICATION NUMBER: 09/247,806
PRIOR FILING DATE: 1999-02-11
PRIOR PELLAGATION NUMBER: FR 98/01614
PRIOR FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SCOFTWARE: PATENTIN VET: 2.1
SEQ ID NO 10
LENGTH: 714
                                                                                                                       TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 G00 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              699 005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                667
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               유
```

RESULT 6 US-09-861-597-2

```
APPLICANT: GARGON, Jean-Claude
APPLICANT: GARGON, Jean-Claude
APPLICANT: ARRAUDEAU, Jean-Pierre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT FILICATION NUMBER: US/09/861,597
CURRENT FILING DATE: 2001-05-22
PRIOR PAPLICATION NUMBER: 09/247,806
PRIOR PAPLICATION NUMBER: PR 98/01614
PRIOR FILING DATE: 1999-02-11
PRIOR FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 2
LENGTH: $29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    264 QQGPSGPG------SAAAAAAAAAAGPGGYG------301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           302 GPSGAGSAAAAAAAAAGPGQQGLGGYG--------PGQQGPG 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        409 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAA-GGAGQGGYGGLGSQGTSGRGGLGGQGA 467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31.4%; Score 1097.5; DB 10; Length 529;
45.3%; Pred. No. 1.1e-56;
Live 22; Mismatches 156; Indels 197; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---IAASAASAGPGGYG-----PAQQGPAGYGPGSAVAASAGA---- 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 GGYGPGQ-------QGPGGYGP-GQQGPSGPG------SAAAAAAAAAAAAGPGGYG 42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     77 AGQGGYGGLG-SQGTSGRGGLGGQGAGAAAAAAAAAGGA--GQGGYG------GLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                103 QQGPSGPG-----SAAAASAAASAESGGPGGYG----PGQQGPGGYGPGGYGPGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               289 SQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         528 AGGAGQGGYGGLGSQGTSGRGGL--GGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     334 GYG------PGQQGPGGYGPGSASAAAAAGPGQQGPGGYGP-GQQGPSGPGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGLGGQGAGAAAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                646 GQGGYGGLGSQGTSGPGGYGPGQQTS 671
Sequence 2, Application US/09861597
Patent No. US200200064539A1
GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 311; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; ORGANISM: Nephila clavipes US-09-861-597-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
```

```
R FILING DATE: 1998-03-20

R APPLICATION NUMBER: 60/079664

R FILING DATE: 1998-03-20

R APPLICATION NUMBER: 60/079664

R FILING DATE: 1998-03-27

R APPLICATION NUMBER: 60/080107

R FILING DATE: 1998-03-31

R FILING DATE: 1998-04-01

R FILING DATE: 1998-04-03

R FILING DATE: 1998-05-05

R FILING DATE: 1998-05-05

R FILING DATE: 1998-05-05

R FILING DATE: 1998-05-05

R FILING DATE: 1998-05-07

R FILING DATE: 1998-05-05

R FILING DATE: 1998-05-15

R FILING DATE: 1998-05-05

R F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R FILING DATE: 1998-05-18
R APPLICATION NUMBER: 60/086392
R FILING DATE: 1998-05-22
R APPLICATION NUMBER: 60/086486
R AFLING DATE: 1998-05-22
R APPLICATION NUMBER: 60/087098
R FILING DATE: 1998-05-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: 60/087208
FILING DATE: 1998-05-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: 60/087609
   APPLICANT: GOGGGACA, AUGTES,
APPLICANT: GOGGGACA, AUGTES,
APPLICANT: GUTNEY, AUGTEN I.
APPLICANT: GUTNEY, AUGTEN I.
APPLICANT: SENITA, VICTORIA
APPLICANT: SENITA, VICTORIA
APPLICANT: SENITA, VICTORIA
APPLICANT: Watanabe, Colin K.
APPLICANT: APADENCA SENITARIA I.
APPLICANT: Zhang, Zemin I.
APPLICANT: Zhang, Zemin I.
APPLICANT: Zhang, Zemin I.
APPLICANT: AROCA, WILLilam I.
APPLICANTON: ACIDS ENCODING THE SAME
TITLE OF INVENTION: ACIDS ENCODING THE SAME
TITLE OF INVENTION: ACIDS ENCODING THE SAME
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: 00/059263
PRIOR FILING DATE: 1997-09-18
PRIOR PELICATION NUMBER: 60/063250
PRIOR FILING DATE: 1997-10-24
PRIOR FILING DATE: 1997-10-24
PRIOR FILING DATE: 1997-10-24
PRIOR FILING DATE: 1997-10-24
PRIOR FILING DATE: 1997-10-21
PRIOR FILING DATE: 1997-10-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  R APPLICATION NUMBER: 60/063544

R FILING DATE: 1997-10-28

R FILING DATE: 1997-10-28

R FILING DATE: 1997-10-28

R APPLICATION NUMBER: 60/063734

R APPLICATION NUMBER: 60/063734

R APPLICATION NUMBER: 60/06370

R FILING DATE: 1997-10-31

R APPLICATION NUMBER: 60/064103

R APPLICATION NUMBER: 60/064103

R APPLICATION NUMBER: 60/0641103

R APPLICATION NUMBER: 60/06311

R APPLICATION NUMBER: 60/065311

R APPLICATION NUMBER: 60/065311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   R APPLICATION NUMBER: 60/066120
R FILING DATE: 1997-11-21
R PILING DATE: 1997-11-24
R APPLICATION NUMBER: 60/06672
R APPLICATION NUMBER: 60/06677
R APPLICATION NUMBER: 60/069335
R APPLICATION NUMBER: 60/069335
R APPLICATION NUMBER: 60/069335
| ||||| | |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: 60/077632
FILLING DAFE: 1998-03-11
APPLICATION NUMBER: 60/077649
FILLING DATE: 1998-03-11
                                                                                                               Sequence 97, Application US/10052586 Patent No. US20020127584A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: 60/069425
FILING DATE: 1997-12-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: 60/069870 FILING DATE: 1997-12-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: 60/068017
FILING DATE: 1997-12-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: 60/077450
FILING DATE: 1998-03-10
                                                                                                                                                                                                                               Goddard, Audrey
Godowski, Paul J.
Gurney, Austin L.
                                                                                                                                                 GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
                                                                                                                                                                                                               Desnoyers, Luc
                                                                                             -10-052-586-97
                 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR
PRIOR
                                                                             RESULT 7
                 g
```

```
1291 TGAAGCTTTGAGGAAGCAACAGTCAAAGAAGGGAACTTCCCGG-GAAGATCAGACCCTTG 1349
                                                                                                                                                                                                                                                                                                                                                                                                                                 1052 AAATTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAG 1111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1112 AAATCAGTCAGCGCAGTGAAGAGCTCAGAAAGAAGCAAGACAATTAAAACGG-GAACT 1170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1350 CACTGCTGAACCAGTTTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAATG 1409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1410 ACATTCCTGAAACAGAAGTAGAAGATGATGAAG-GATGGATGTCACATGTACTTCAGTTT 1468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1469 GAGGATAAAAGCAGAAAAGTGAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATC 1528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1529 TATGATCCTCGGAATCCAGTGAATAAAGAAGGAGGAAGAAAGCAAAAAGCTGATGAGA 1588
                                                                                                                   819 AAGCTGAGGAAGAAGGAAGGAAGTAAATCGAGTTAGTCAGAGCATGAAGGGAAAAGCA 878
                                                                                                                                                                                                                                                                          939 GTGAAAAAGGTGATGCACCAGATTTAGTTGATGAGGAAGATGAAAGTGCAGAGCATG 998
                                                                                                                                                                                                                                                                                                                                                                                           ---GQGAGAA 192
                                                                                                                                                                                                                                                                                                                    381 SGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLG-----SQGTSGRGGLGGQGAGAAA 433
                                               759 AAGTAAAGAAATTGAAACCCAAAAGGCACAAAAATTTTAGTTTACTTTCATTTGGAGAGG
                                                                                                                                                                                                                                        193 AAAAAAAGGGGG----YGGLGSQGTSG----RGGLGGQGAGAAAAAAAAGGAGQG
                                                                                                                                                                                                    879 AAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGTTGTAGAAA
                                                                                                                                                                                                                                                                                                                                                       434 AAAAAAAGGAG------QGGYGGLGSQGTSGRGG-----LGGQGAGAAAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               518 -----GAAAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         338 P------GQQTSGRGGLGGQGAGAAAAAAAAGGA-----GQGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----LGG-----QGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGA
                                                                                 120 YGG-LGSQGTSGRGGLGGQGAGAAAAAAAAAG------GAGQGGYGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   482 ------GQGGYGG----LGSQGTSGPGGYGPGQQTSGRGGLGGQGA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION: PRODUCTION OF CHIMERIC CAPSID VECTORS FILE REFERENCE: 102182-14
CURRENT APPLICATION NUMBER: US/09/804,898
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/189,110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1589 GAGAAAAAAAAA -----GAAGATAAATGAGAATAATGATAA 1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----AAAAAAGGAG-QGGYGGLGSQGTSGRGGLGG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  554 GAGAAAAAAAAGGAGGGYGGLGSQG-------
             -----LGSQGTSGRGGLGGQGAGAAAAAAAA
                                                                                                                                                              --LGSQGTSGPGGYGPGQQTSGRGGLG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
US-09-804-898-2
US-09-804-898-2
Sequence 2, Application US/09804898
Patent No. US20020045264A1
GENERAL INFORMATION:
APPLICANT: DURING, MATTHEW
APPLICANT: XIAO, WEIDONG
                                                                                                                                                              161
           85
                                             셤
                                                                                     ò
                                                                                                                       g
                                                                                                                                                              ŏ
                                                                                                                                                                                                g
                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         44 -------YGGLGSQGTSGRGGLGGQGAAAAAAAAAAAGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        357;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1894;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R FILING DATE: 1998-06-02
R PAPLICATION NUMBER: 60/087759
R PILING DATE: 1998-06-02
R APPLICATION NUMBER: 60/087827
R FILING DATE: 1998-06-03
R PILING DATE: 1998-06-04
R APPLICATION NUMBER: 60/088028
R FILING DATE: 1998-06-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                           R FILING DATE: 1998-06-09
R APPLICATION NUMBER: 60/080722
RR APPLICATION NUMBER: 60/080738
R FILING DATE: 1998-06-10
R FILING DATE: 1998-06-10
R FILING DATE: 1998-06-10
R FILING DATE: 1998-06-10
R APPLICATION NUMBER: 60/080811
R APPLICATION NUMBER: 60/080811
R APPLICATION NUMBER: 60/080818
R APPLICATION NUMBER: 60/080825
R FILING DATE: 1998-06-10
R APPLICATION NUMBER: 60/080825
R FILING DATE: 1998-06-10
R APPLICATION NUMBER: 60/080825
R FILING DATE: 1998-06-10
                                                                                                                                                                                                                                                                    R FILING DATE: 1998-06-05
R APPLICATION NUMBER: 60/086202
R FILING DATE: 1998-06-05
R APPLICATION NUMBER: 60/086212
R FILING DATE: 1998-06-05
R APPLICATION NUMBER: 60/086217
R APPLICATION NUMBER: 60/086326
R APPLICATION NUMBER: 60/08635
R APPLICATION NUMBER: 60/08655
R APPLICATION NUMBER: 60/08655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RELIENG DATE: 1998-06-10
RAPELICATION NUMBER: 60/08861
RELING DATE: 1998-06-10
RAPELICATION NUMBER: 60/08863
RFILING DATE: 1998-06-11
RAPELICATION NUMBER: 60/08876
RFILING DATE: 1998-06-11
RAPELICATION NUMBER: 60/08976
RAPELICATION NUMBER: 60/089105
RAPELICATION NUMBER: 60/089105
RELING DATE: 1998-06-12
RELING DATE: 1998-06-16
RAPELICATION NUMBER: 60/089518
RELING DATE: 1998-06-16
RAPELICATION NUMBER: 60/089518
RELING DATE: 1998-06-16
RELING DATE: 1998-06-17
                                                                                                                                                                                                                                  FILING DATE: 1998-06-04
APPLICATION NUMBER: 60/088167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION NUMBER: 60/089908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 30.99
Matches 311; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                           PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

ð

```
g
    à
                                                                                                                                                                                                                     36;
                                                                                                                                                                                                                                                                                                                                   56 GGLGGQGAGAAAAAAAAAAAGG----AGQGGYGGLGSQCTSGRGGLG-----GQGAG----- 102
                                                                                                                                                                                                                                                                                                                                                           162 -GACGTGAATTACGTCATAGGGTTAG-GGAGGTCCTGTATTAGAGGTCACGTGAGTTT 219
                                                                                                                                                                                                                                                                                                                                                                                                                                            220 TGCGACATTTGCGACACCATGTGGTCACGCTGGGTATTTAAGCCCGAGTGAGCACGCAG 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         270 AAAAAAAGGAGGG-----GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAG 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 519 TGGCGCCGTGTGAGGCCCCCGGAGGCCCTTTTCTTTGTGCAATTTGAGAAGGGAGAG 578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       355 A----GLGSQGTSGRGG 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              579 AGCTACTICCACATGCACGTGCTCGTGGAAACCACGGGGTGAAATCCATGGTTTTGGGA 638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      699 ACTITGCCAAACTGGTTCGCGGTCACAAAGACCAGAAATGGCGCCGGAGGGGGGAACAAG 758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GTSGRGGLGGQGAGAAAA-----AG 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GIGGIGGAIGAGIGCIACAICCCCAATIACTIGCICCCCAAAACCCAGGCICGAG 818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GAGQGGYGGLGSQGTSGPGQQTSGRGGLG---GQGAGAA-----AAAAAAAA 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      879 CGGTIGGIGGCGCAGCAICIGACGCACGIGICGCAGACGCAGGAGCAGAACAAAGAGAAI 938
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               574 YGGLGS------AGTSGRGGLGGQG---AGAAAAAAAA-------AGAGAG 608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         145 ------PGGYGPGQGTSG---QGGYGGLGSQGTSG-----PGGYGPGQQTSGRGG 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                280 GGTCTCCATTTTGAAGCGGGAGGTTTGAACGCGCCAGCCGCCATGCCGGGGTTTTACGAGA 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             340 TIGIGATIAAGGICCCCAGCGACCTIGACGAGCATCTGCCCGGCATITCTGACAGCTITG 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        459 CTGATTGAGCAGGCACCCCTGACCGTGGCCGAGAAGCTGCAGCGCGACTTCTGACGGAA 518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     386 LG------GQGAGAAAAAAAAAAAAAAGGAGQG------GYGGLGSQG----- 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         639 CGTTTCCTGAGTCAGATTCGCGAAAACTGATTCAGAGAATTTACCGCGGGATCGAGCCG 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             819 TGGGCGTGGACTAATATGGAACAGTATTTAAGCGCCTGTTTGAATCTCACGGAGCGTAAA 878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGAGQGGYGGLG------SQGTSGRGCLGGQGAGAAAAAAAAAAAGGAGQGG 573
                                                                                                                                                                                                                                                                                103 -----AAAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAA---- 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QGGYGGLGS---QGTSGPG------GYGPGQQTSGRGGL-GGQGAG 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----TSGRGCLGGQGAGAAAAAAAAGGAG-QGGYGGLG----SQ 454
                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                        11 GSMASGRGGLGGQGAGAAAAAAAAAAA------GGAGQGGYGGLGSQGTSGR 55
                                                                                                                                                                          21.1%; Score 738; DB 10; Length 4679; 30.0%; Pred. No. 1.5e-35; ive 21; Mismatches 338; Indels 312;
                                                                                               ; TYPE: PRT
; ORGANISM: adeno-associated virus 2
US-09-804-898-2
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin version 3.0
                                                                                                                                                                        Query Match 21.1%
Best Local Similarity 30.0%
Matches 288; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------
                                                                             LENGTH: 4679
                                                          SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      480
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
```

R FILING DATE: 1998-03-20
R APPLICATION NUMBER: 60/078939
R APPLICATION NUMBER: 60/079664
R FILING DATE: 1998-03-27
R APPLICATION NUMBER: 60/079786
R FILING DATE: 1998-03-27
R FILING DATE: 1998-03-27
R FILING DATE: 1998-01

FILING DATE: 1998-03-31 APPLICATION NUMBER: 60/080194 FILING DATE: 1998-03-31

FILING DATE: 1998-03 APPLICATION NUMBER:

FILING DATE: 1998-03-11 APPLICATION NUMBER: 60/078886

FILING DATE: 1998-04-01 APPLICATION NUMBER: 60/081049 FILING DATE: 1998-04-08

PRIOR

FILING DATE: 1998-04-01 APPLICATION NUMBER: 60/080333

60/080327

APPLICATION NUMBER: 60/081070 FILING DATE: 1998-04-08 APPLICATION NUMBER: 60/081195 FILING DATE: 1998-04-09

APPLICATION NUMBER: 60/081838 FILING DATE: 1998-04-15 APPLICATION NUMBER: 60/082568

```
Query Match 19.6%; Score 686; DB 12; Length 1300;
Best Local Similarity 28.1%; Pred. No. 5.7e-33;
Matches 277; Conservative 16; Mismatches 347; Indels 346; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        182 ACGAGACGGAGGAGGTCCCAAAGATGCCCTCCGAGCAGTAAAGAAGAGAATCGTGGGGA 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64 GAAAAAAAAAAAGGAGGGYGGLGSQG------TSGRGG-LGGQGAGAAAAA 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       242 ATAAGAACTTCCACGAGGTGATGCTGGCTCTCACAGTCTTAGAAACCTGTGTCAAGAACT 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14 ASGRGGLGGQGAGAAAAAAAAAAGGAG-----QGGYGGLGSQGTSGRG----GLGGQGA 63
                                                                          R APPLICATION NUMBER: 60/087827
R FILING DATE: 1998-06-03
R FILING DATE: 1998-06-04
R FILING DATE: 1998-06-04
R FILING DATE: 1998-06-04
R FILING DATE: 1998-06-04
R APPLICATION NUMBER: 60/088028
R APPLICATION NUMBER: 60/088028
                                                                                                                                                                                                                                            APPLICATION NUMBER: 60/088033
FILING DATE: 1998-06-04
APPLICATION NUMBER: 60/088167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 1998-06-04
APPLICATION NUMBER: 60/088655
60/08/609
                                                                                                                                                                                                                                                                                                                                60/088202
                                                                                                                                                                                                                                                                                                                                                                        60/088212
                                                                                                                                                                                                                                                                                                                                                                                                                                                       60/088326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: 60/088722
FILING DATE: 1998-06-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 60/088738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     60/088740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              60/088811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: 60/088824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60/088825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: 60/088826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: 60/088861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: 60/088863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60/088876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     060680/09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60/089512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: 60/089653
FILING DATE: 1998-06-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: 60/089908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 1998-06-10 APPLICATION NUMBER: 60/0
                                                         1998-06-02
                                                                                                                                                                                                                                                                                                                                                   1998-06-05
                                                                                                                                                                                                                                                                                                                                                                                           1998-06-05
                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: 1998-06-05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 1998-06-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 1998-06-10
                                                                                                                                                                                                                             FILING DATE: 1998-06-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: 1998-06-09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1998-06-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 1998-06-16 APPLICATION NUMBER: 60/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91-90-8661
                                                                                                                                                                                                                                                                                                        FILING DATE: 1998-06-05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE: 1998-06-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 998-06-17
                   1998-06-0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: 1998-06-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE: 1998-06-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE: 1998-06-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: 1998-06-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 1998-06-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -90-866
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR FILING DATE: ]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR
                                                                                                                                                                                                                                                                                                                                                                                         PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR
PRIOR
PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

R FILING DATE: 1998-04-21 R FILING DATE: 1998-04-21 R FILING DATE: 1998-04-21 R APPLICATION NUMBER: 60/082704 R APPLICATION NUMBER: 60/082797 R APPLICATION NUMBER: 60/082797 R FILING DATE: 1998-04-22

FILING DATE: 1998-04-29
APPLICATION UNDBER: 60/083499
FILING DATE: 1998-04-29
APPLICATION NUMBER: 60/083559

APPLICATION NUMBER: 60/083495 FILING DATE: 1998-04-29

APPLICATION NUMBER: 60/083496

NPPLICATION NUMBER: 60/083322 FILING DATE: 1998-04-28

FILING DATE: 1998-05-05 APPLICATION NUMBER: 60/084414 FILING DATE: 1998-05-06

APPLICATION NUMBER: 60/084639 FILING DATE: 1998-05-07 APPLICATION NUMBER: 60/084640 APPLICATION NUMBER: 60/084643

1998-05-07

ILING DATE:

FILING DATE: 1998-04-29 APPLICATION NUMBER: 60/084366

38;

R FILING DATE: 1998-05-15
R APPLICATION NUMBER: 60/085700
R FILING DATE: 1998-05-15
R APPLICATION NUMBER: 60/086023
R FILING DATE: 1998-05-18
R APPLICATION NUMBER: 60/086392

FILING DATE: 1998-05-22 APPLICATION NUMBER: 60/086486

FILING DATE: 1998-05-22

APPLICATION NUMBER: 60/087098 FILING DATE: 1998-05-28 APPLICATION NUMBER: 60/087208 FILING DATE: 1998-05-28

APPLICATION NUMBER: 60/085579 FILLING DATE: 1998-05-15 APPLICATION NUMBER: 60/08580 APPLICATION NUMBER: 60/085582

FILING DATE: 1998-05-07 APPLICATION NUMBER: 60/085573 FILING DATE: 1998-05-15

PRIOR

σ

28;

Gaps

```
GACGGTTCCGGCCAATACAGCTGGATCATTAACGGAATCGAGTGGGCGATCGCAACAAT 770
                                                                                                                                                                                                                                                                                                 125 GCTGTTTGCTTTAGCGTTAATCTTTACGATGGCGTTCGGCAGCACATCCTCTGCCCAGGC 184
                                                                                                                                                                                                                                                                                                                                                                                                                 CATCCTGATTTAAAGGTAGCAAGCGGAGCCAGCATGGTTCCTTCTGAAACAAATCCTTTC 590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     70 AAATGAAAAAAGGAGAG-----GATAAAGAGTGAGAGGCAAAAAAGTATGGATCAGTTT 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SGRGCLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA--GAAAAAAA,564
                                                                                                                                                                                                                                                                                                                               -----RGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GTTGATAAAGCCGTTGCATCCGGCGTCGTAGTCGTTGCGGCAGCCGGTAACGAAGGCACT
                                                                                                                                                                                                                                                                                                                                                                                               155 QGGYG-GLGSQGTSGPGGYGPGQQTSG-----RGGLGGQGAGAAAAAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                           GQGGYGGLGSQGTSGRGGLGG--QGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        262 GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                471 CAAGGCTACACTGGATCAATGTTAAAGTAGCGGTTATCGACAGCGGTATCGATTCTTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----QGGYGGLGSQGTSG-RGGLGGQG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --AGAAAAAA----AAAGGAG------QGGYGGLGSQGTSGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATCGGTGTATTAGGCGTTGCGCCAAGCGCATCACTTTACGCTGTAAAAGTTCTCGGTGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATGGACGTTATTAACATGAGCCTCGGCGGACCTTCTGGTTCTGCTGCTTTAAAAGCGGCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   358 AATTGAAAAAGACCCGAGCGTCGCTTACGTTGAAGAAG---ATCACGTAGCACATGCGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      322 YGGLGSQGTSGPGGYGPGQQTSGRGCLGGQGAGAAAAAAAAA---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   363 -AAGG----AGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAG-----3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAAGGAG----QGGYG-----YGPGQQT
                                                                                                                                           Length 1497;
                                                                                                                                                                          Indels 245;
                                                                                                                                                                                                      DB 10;
                                                                                                                                         Query Match 19.3%; Score 675.5; DB 10; Best Local Similarity 30.7%; Pred. No. 2.5e-32; Matches 265; Conservative 22; Mismatches 330;
                                3.0
CURRENT FILING DATE: 1998-04-15
WHMER OF SEQ ID NOS: 7
SOFTWARE: FastSEQ for Windows Version
SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----AGQGGYGGLGSQGTSG 660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             891 TCCGGCAGCTCAAGCACAGTGG 912
                                                                                           ; ORGANISM: B. amyloliquefaciens US-09-060-854B-2
                                                            LENGTH: 1497
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         651
                                                                                                                                                                                                                                                                    7.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                             204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    771
                                                                                                                                                                                                                                                                                                                                97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       201
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/09060854B
Patent No. US20020081703A1
GENERAL INFORMATION:
APPLICANT: Estell, David Aaron
TITLE OF INVENTION: Human Protease and Use of Such Protease for Pharmaceutical
TITLE OF INVENTION: Applications and for Reducing the Allergenicity of No. US2002008
TITLE OF INVENTION: Proteins
FILE REFERENCE: GC532
CURRENT APPLICATION NUMBER: US/09/060,854B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TGGCCACTGAGGGTAGGGTGTGGAGGCCCCCTGAGGAGCTGCGCGGCCCAG 1076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             958 ATTCTGCCGAGCCTATTGTCCCTACCCTGCTCTGCTGCATGGGGCCCCATGGCTTTG-GC 1016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTACGAAGCTGCAACTCTGCGCGCAGTG----GGCGAGATCTCATCAGCCCCAGGCTGC 1131
                                                                                                                                         206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      391 AGAAAAAAAAAAAGG------AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA 440
              AAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGP 169
                                                                                                                                                                       481
                                                                                                                                                                                                      254
                                                                                                                                                                                                                                     541
                                                                                                                                                                                                                                                                    284
                                                                                                                                                                                                                                                                                                 601
                                                                                                                                                                                                                                                                                                                                297
                                                                                                                                                                                                                                                                                                                                                                                               334
                                                                                                                                                                                                                                                                                                                                                                                                                            721
                                                                                                                                                                                                                                                                                                                                                                                                                                                          335 GYGPGQQ-TSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSG-RGGLGGQG-- 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     550 LGG----QGAGAAAAAAAAAAGGAGQGGYGGLGS-----QGTSGRGGLGGQGAGAAAAA 600
                                    302 GCGGGCACCGCTTCCACGTGCTGGTCGTCGACGACTTCGTGGAGAGTGTGGTGA 361
                                                                                                        362 GGACCATCCTGCCCAAGAACAACCACCACCATCGTGCATGACAAAGTGCTCAACGTCA 421
                                                                                                                                                                                                                                                                                                                                                                661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGAGQGGYG--GLGSQGTSGRGG---LGGQGAGA-------AAAAAAAA 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                838 cerecceaeccarecaecaecaeceserereaerearaecerecreesecceareceaa 897
                                                                                                                                                                                                                                                                                      662 GCCCATACTCTCCGGTGACACGCCCATAGCACCACAGCCCGGAACAGATTGGGAAGCTGCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CAGTGAGCTGGAGATGGTGGGAACGTGAGGGTGATGTCGGAGATGCTGACGGAGCT
                                                                                                                                                                                                                                 -----SGRGG-LGGQG-----AGA-AAAAAAAAGGAGQGGY-----
                                                                                                                                                                                                                                                                                                                                                              522 AAAAAAGGAG-----GLGSQGTSGRGG
                                                                                                                                                                     422 TCCAGTCCTGGGCTGACGCGTTCCGCAGCTCGCCCGATCTGACAGGTGTGGTCACCATCT
                                                                                                                                                                                                    --GYGGLGSQGTSGRGG-----LGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT---
                                                                                                                                                                                                                                                                                                                                                                                               ------LGGQG-----AGAAAAAAAAAAGGAG-QGGYGGLGSQGTSGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  782 GGTGCCCACCCAGGCCCGAGCCCGCAGCTGCTGCTGCAG----GAGCTCAACCGCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       601 AAAAGGAGQGGYG-----GLGSQGTSGRGGLGGQGAGAAAAAAAAAAAA
                                                                                                                                       -----SGRGGLGGQGAG------AAAAAAAAAAAGGAGQG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGAG--QGGYGGLG-----QGAGAAPGQYGPGQQTSGRGGLGG----QGAGAAA
                                                                                                                                                                                                                                                                                                                                -GGLGSQG-----TSGRGG-----TSGRGG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AGGTGAGGCTTCAGGGGATGCTGGGG 1157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AGGAGQGG----YGGLGSQGTSGPGG 663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
US-09-060-854B-2
              110
                                                                         170
                                                                                                                                      179
                                                                                                                                                                                                    207
                                                                                                                                                                                                                                                                  255
                                                                                                                                                                                                                                                                                                542
                                                                                                                                                                                                                                                                                                                                285
                                                                                                                                                                                                                                                                                                                                                              602
                                                                                                                                                                                                                                                                                                                                                                                           298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1132
                                       g
                                                                         οy
                                                                                                     g
                                                                                                                                    ò
                                                                                                                                                              셤
                                                                                                                                                                                               à
                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
```

470

261 357 321 414 428

```
US-10-096-961-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     270
                                                                                   q
                                                                                                                                                                                                                             qq
                     ö
                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             33;
                                                                                                                                                                                                       TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO A CARDIAC-SPECIFIC TITLE OF INVENTION: MUCLEAR REGULATORY FACTOR FILE REPRENCE: UTSD.6950.8 CURRENT APPLICATION NUMBER: US/10/029,217A CURRENT APPLICATION NUMBER: US/10/029,217A PRIOR APPLICATION NUMBER: 60/257,761 PRIOR PILIG DATE: 2000-12-21 PRIOR FILING DATE: 2000-12-21 SUPPRIOR FILING DATE: 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1004 AACAGCACCTCACTGACTGGCAAGCCG--GGAGCCCTGCCGGCCAACCTGGACGACATGA 1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1062 AGGTGGCAGAGCTGAAGCAGGAGCTGAAGTTGCGATCACTGCCTGTCTCGGGCACCAAAA 1121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1122 CTGAGCTGATTGAGCGCCTTCGAGCCTATCAAGACCAAATCAGCCCTGTGCCAGGAGCCC 1181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1182 ccaasicccricciccaccritaricaricaria actionaria estas est
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     240 CCCAGCGGCCCGGCTGAGCACGGGGCCAGCCCTGGTGGCAGCAGCAGGCTTGGCTCCAGCTGA 1299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1353 GCTCCACGCCCCCGTGTCTCCCACCCCTCGGAGGGTTCACTGCTCACACGGGGGGATG 1412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1473 CCCTGCAGGCCTCGCCACTGCAGATCCTCGTGAAGGAGGAGGCCCCCGGGCCGGGTCCT 1532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --AAAAAAAAAAAGGAG------QGGY-----GGLGSQGTSGRGGLGGQGAGAAA 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155 QGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGA-------GAAAAAAAA 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                200 AGGAGQGGY---GGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGG---YGGLGSQG 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          429 AGAA-----RGELGGQGAGAA 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               886 GCCAAAGTCAGCAGGCGAGGCCTGG-GAAGCAGCGGGACCCCCCCAGTACGCAGCCTCT 944
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24 GAGAAAAAAAAAAGGAGQGG---YGGIJGSQGTSGRGG------LGGQGAG----- 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                295 --RGGLGG-----QGAGAAAAAAAAAGGAGGGGYGGLGS------QGTSGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      334 GGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1413 AAAACTCCACCCCGGGGACACCTTTGGTGAGATGGTGACATCACCTCTGACGCAGCTGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                393 -----GYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            471 AAAAAAAAGGG-----GYGG----LGSQGTSGPGGYGPGQQTSGRGGLGGQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 19.2%; Score 672; DB 9; Length 3907; Best Local Similarity 30.9%; Pred. No. 7.7e-32; Matches 258; Conservative 15; Mismatches 342; Indels 220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 24, Application US/10029217A Patent No. US20020164735A1
                                                                                                     GENERAL INFORMATION:
APPLICANT: OLSON, ERIC N.
APPLICANT: WANG, DA-ZHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; ORGANISM: Mus musculus
US-10-029-217A-24
US-10-029-217A-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
WS-10-096-961-1

Sequence 1, Application US/10096961

Sequence 1, Application US/10096961

Sequence 1, Application US/10096961

GENERAL INFORMATION:

APPLICANT: GUEGLER, Karl et al.

TITLE OF INVENTION: ISCLATED HUMAN PHOSPHOLIPASE PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE

TITLE OF INVENTION: ROCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE

TITLE OF INVENTION: ROCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE

TITLE OF INVENTION: NUMBER: US/10/096,961

CURRENT FILING DATE: 2000-09-14

PRIOR FILING DATE: 2000-09-14

PRIOR FILING DATE: 2000-09-14

PRIOR FILING DATE: 2000-12-18

NUMBER: OF SEQ ID NOS: 5

SOCTWARE: FASTESEQ for Windows Version 4.0

LENGTH: 2211

LENGTH: 2211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         37;
1533 GTTGCCTGAGCCCTGGGGGGGGGGGGGGGGTAGAGG-GGCGCGACAAGGACCAGATGCTG 1591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          57 GLGGQGAGAAAAAAAAAAGGAG-QGGYGGLGS-----QGTSGRGGLGGQGAGAAA 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 544 GCGTGCTGAGTGCCCCTGAGCTGCTGGAGAGAGGAGGGCGAGGAGAG---GGC 699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     525 GGAGTGTGACCACTCCCAACAACGACCGTCTAGAGGGGGCCTGAGATCGAGGAGTTCCTGCG 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----- GGGGAGAAAA 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           292 TSGRG--GLGG------QGAGAAAAAAAAAAAGGAGG------GYGGLGSQGTSGPG 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      335 GYGPGQQTSG--RGGLGGQGAG----AAAAAAAAAAAAGG------AGQG 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                936 CAGCACCGAGGCCTATGTTAGGGCCTTTGCCCAGGGATGCCGCTGCGTGGAGCTGGACTG 995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 309; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19 GLGGQGAGAAAA----AAAAAAG-----GAGQG------GYGGLGSQGTSGRG 56
                                                                        574 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGA--GQGGYGGLG-----SQGTSGRGG
                                                                                                                                      1592 CAG------GAGAAAGACAAGCAGATCGAGGCGCTGACGCGCATGCTCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 107 AAAAAAGGAGGG-----GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             700 GCCACACTGGCCCGCGCCCAGCTCATTCAGACCTATGAGCTCAACGAGACAGCCAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  196 AAAAAGGAG-----OGGYGG-----CGSQGTSG-RGGLGGQGA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      820 GCCITGACAACACCCACACACT ---- GTGTGTTCCAGGACATGAACCAGCCCCTTGCCCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------AGGAGOGGYGCLGSOG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         876 CTACTICATCITCTICCTCCCACACACCTATCIGACTGACTCCCAGATCGGGGGGCCCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 228 -----GAAAAAAAAAAGGGGGGGGGGGGGGGGGGGAA-----
                                                                                                                                                                                                                                                                                     1637 CAGAAGCAGCAGCTGGAGCGGCTCAAGCTGCAGCTGGAGCAGGAGAAGCGAG 1691
                                                                                                                                                                                                               626 LGGQGAGAAAAAAAAAAGGAGQGG------YGGLGSQGTSGPGGYGPG 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 2211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 18.8%; Score 656.5; DB 9;
Best Local Similarity 29.2%; Pred. No. 3.9e-31;
Matches 274; Conservative 17; Mismatches 337;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               162 GSQGTSGPGG--YGPGQQTSGRGGL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Homo sapien
```

```
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                   119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          297
                                                                                                                                                                                                                                                                                                                                                                                                                                       185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       512
                                                                                                                                                                                                ò
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                 á
                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MCCONNELL, Stephen J.
INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME
                                                                                                                                                                                                                                                                                             1294 CCCGCTGCTCGGAGCGAGGATG--GCCG-GGCTCTGTCGGATCGGGAGGAGGAGGAG 1350
CTGGGAGGGCCAGGAGGGGAGCCCGTCATCTATCATGGCCATACCCTCACCTCCAAGAT 1055
                                                       1056 TCTCTTCCGGGACGTGGCCCAAGCCGTGCGAGCCATGCCTTCACGCTGTCCCCTTACCC 1115
                                                                                                                  1116 TGTCATCCTATCCCTGGAGAACCACTGGGGGCTGGAGCAGCA--GGCTGCCATGGCCCGC 1173
                                                                                                                                                                             1174 CACCTCTGCACCATCCTGGGGGACATGCTGGTGACACAGGCGCTGGACTCCCCAAATCCC 1233
                                                                                                                                                                                                                            ----AAAAGGAGQGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAGGAGQGGYGGLG 578
                                                                                                                                                                                                                                                                                                                             SQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 638
                                                                                                                                                                                                           477 AAGGAGGG-----YGGLGSQGTSGPGGYG-PGQQTSGRGGLGGQGAGAAAAA----
                           -----GYGGLGSQGTSGRGGLGGQGAG------
                                                                                        STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: CURNOWN>
                                                                                                                                                                                                                                                                                                                                                                                                                    AGATCTCCCCGGAGCTGTCGGCCCTGGCTGTACTG 1445
                                                                                                                                                                                                                                                                                                                                                                                        AAA-----AGGAGQGG-----YGGLGSQGTSGPG 662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORNEY/AGENTINEORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEFAX: (212) 790-9090
TELEFAX: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 37:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: SPARKS, Andrew B.
HOFFMAN, No. US20020034755A1h
KAY, Brian K.
FOWLKES, Dana M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION DATA:
APPLICATION NUMBER: US 08/630,915
FILING DATE: 03-APR-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 37, Application US/09879957 Patent No. US20020034755A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 22 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
US-09-879-957-37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR
                                                                                                                                                                                                                                                                    525
                                                                                      393
                                                                                                                                                430
                                                                                                                                                                                                                                                                                                                                                           1351
                                                                                                                                                                                                                                                                                                                                                                                          639
                                                                                                                                                                                                                                                                                                                                                                                                                    1409
                                                      셤
                                                                                                                  셤
                                                                                                                                                                         qq
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                     δý
                                                                                                                                                ò
                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                   ò
```

```
32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                245 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRG 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   339 ATGCAGAGAAAATCCCAGAAAATGAGTTCCCGCTCCAGTGAAACCAGTGACTGATTCAA 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AGGAGGGGYGGLGSQGTSGRGGLGG------QGAGAAAAAAAAAAAGGACQG 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGAAGCACAAGACAAGCTGGGTCGGCTTTTCCATCAACACCAAGAACCAGCTAAGCCA-- 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               224 GATGAAATCACTATCCAGCCAGGAGACATAGTCATGGTGGATGAAAGCCAAACTGGAGAA 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 399 CATCTGCCCCTGCCCCAAACTGGCCTTGCGTGAGCCCCCGCCCCTTTGGCAGTAACCT 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  459 CTTCAGAGCCCTCCACGACCCCTAATAACTGGGCCGACTTCAGCTCCACGTGGCCCACCA 518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------OGAGAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGG---QGAGAAAAAAA 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       638 GGCTCCTCCCGTCTCCTGTGCTAGGCCAG - GGTGAAAAGGTGGAGGGGCTACAAGCTC 695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         696 AAGCCCTATATCCTTGGAGAGCCAAAAAGACAACCACTTAAATTTTAACAAAATGATG 755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   756 TCATCACCGTCCTGGAACAGCAAGACATGTGGTGGTTTTGGAGAAGTTCAAGGTCAGAAGG 815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  572 GGYGGLG----SQGT----SGRGT-GGGGAGAAAAAAAAAAAAGGAG 608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        B16 GITGGITCCCCAAGICTTACGIGAAACTCATITCAGGCCCCATAAGGAAGTCTACAAGCA 875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----AAAAAGGGGGYGGLGSQGTSGPGG------YGPGQQTSGRGG---L 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAG----GAGQGGYGGLGSQGTSGRGGLGGQG------AGAAAAAAAA-- 477
                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                 14 ASGRGCLGGQGAGAAAAAAAAAAAAAAAAAAAAAAAAAA 73
                                                                                                                                                                                                                                                                                                                                                                                                                      2 AAAGGGAGGAGAGTGTCAAAAA--GAAGGATGGCGA------GGAAAAAGGCAAACA 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGQGAGAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGACQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CCCG-GCTGGCTTGGAGGAATTAAAAGGAAAGACAGG----GTGGTTCCCTGCAAACT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQCAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        357 AA------QGT-----QGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          519 GCACGAATGAGAAACCAGAAACGGATAACTGGGATGCATGGGCAGCCCAGCCCTCTCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------GGAGQGGYGGLGSQGTSGRGGLGGQGAGA---AAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----SGRG-------GLGG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----AGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGG---
                                                                                                                                                                                                                                                                                                        Indels 344;
                                                                                                                                                                                                                                            Length 1400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GYGGLGSQGTSGRGG----LGGQ-----GAGAAAA-----
                                                                                                                                                                                                                                         18.7%; Score 652; DB 10;
28.3%; Pred. No. 5.2e-31;
tive 24; Mismatches 325;
               TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-879-957-37
LENGTH: 1400 amino acids
                                                                                                                                                                                                                                         Query Match
Best Local Similarity 28.3%
Matches 274; Conservative
```

```
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ITTLE OF INVENTION: ACIDS ENCODING THE SAME FILE REFERENCE: P3430R1C1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CURRENT APPLICATION NUMBER: US/10/052,586
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR FILING DATE: 1997-09-18
PRIOR PILING DATE: 1997-09-18
PRIOR PLING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/05250
PRIOR PILING DATE: 1997-10-17
PRIOR PILING DATE: 1997-10-17
PRIOR PILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR APPLICATION NUMBER: 60/063121
PRIOR PELING DATE: 1997-10-24
PRIOR FILING DATE: 1997-10-24
PRIOR FILING DATE: 1997-10-24
PRIOR FILING DATE: 1997-10-21
PRIOR FILING DATE: 1997-10-21
PRIOR FILING DATE: 1997-10-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 525, Application US/10052586 Patent No. US20020127584A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Watanabe, Colin K. Wood, William I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Godowski, Paul J. Gurney, Austin L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Smith, Victoria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Goddard, Audrey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Desnoyers, Luc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Baker, Kevin P. APPLICANT: Chen, Jian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Zhanq, Zemin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-10-052-586-525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT
                                                                                            278
                                                                                                                                                                                                                                                                                                                                                                                                          487
                                                                                                                                                                                                 327
                                                                                                                                                                                                                                                                                                        447
                                                                                                                                                                                                                                                                                                                                                                          497
                                                                                                                                                                                                                                                                                                                                                                                                                                              531
                                                                                            ò
                                                                                                                       g
                                                                                                                                                               δ
                                                                                                                                                                                               g
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        35;
609 QGG----YGGLGSQGTSGRGG---LGCQGAGAAAAAAAAAGGAG-----QGGYGGLGS 655
                              876 TGGATTCTGGTTCTTCAGAGAGTCCTAGTCTAAAGCGAGTAGCCACTCTCCCAGCAGCCA 935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             74 AGGAGQGGYGGLGSQGTSGRGG---LGGQGAGAA-----AAAAAAAGGAGQGGYGG 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LGSQCTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGS-----QGTSGPG----- 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      163 CGACTGACTTGTCTTCTTGATG----TCCTTTAAGCCGGA------GCTGATTC 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14 ASGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 187;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18.7%; Score 651.5; DB 10; Length 720; 32.5%; Pred. No. 3.5e-31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Tryggvason, Karl
Kallunki, Pekka
Pyke, Charles
TITLE OF INVENTON: Laminin Chains: Diagnostic Uses
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: Fay Sharpe Fagan Minnich & McKee STREET: 1100 Superior Ave, Suite 700 CITY: Cleveland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      314;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/663,147
FILING DATE: 150-September 2000
ATTORNEY AGENT INFORMATION:
NAME: Minnich, Richard, J.
REGISTRATION NUMBER: 24,175
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 216-861-5582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      17; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-756-071B-20
                                                                                                                                                                                            Sequence 20, Application US/09756071B Patent No. US20020052307Al GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 720 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO: 20 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                        STATE: Ohio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                    656 QGTSGPG 662
                                                                                                         936 AGCCGAG 942
                                                                                                                                                                           US-09-756-071B-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Simi
Matches 249;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 171
 8
                                  셤
                                                                      à
                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
376 TCTGTGCCACACCCTTGGCCCG--GGCCAGGT-----GTGCGCCCTCCTCGCTGCGAG 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GYGGLG-SQGTSGRGGLGGQGAGAAAAAAAAG---GAGQGGYG----GLGSQGTSG-- 496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TGCCTCTGGAGGGCTGGGTCCTCCTTATTCACAGGTGAGTCACACCCTGAAACACAGGCT 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            547 CICITCCIGICAG-GACIGAGICAGGIAGAAGAGCGAIAAAACCACCIGAICAAGGAAA 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 572 --GGYGGLGSQGTSGRG-GLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         606 AGGAAGGCACAGGGAGCGCAGAGTGAGAACCACCAACCGAGGCGCGG------GGCAG 658
                                                                                                                                                                                                                        268 GGGCG-GCTGCTGTCTGTGAATCTGCCCTGGACCACCCCGGGAGAGGAGGAGG
223 GGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA----AAAAAG
                                                                        208 GGGCTGCTGTTTTTTGAGTTAGCGCTCTTAAGATTGGGCCTCCCAGTTTGAGAAG
                                                                                                                                                    GACQGGYGGLGSQGT----SGRGGLGGQG---AGAAAAAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                    -------GCTCCGGGGAATCTCGACATTCCAGGCAAAGGCTCCCGGGCCGCAGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                            387 GGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----PGGYGPGQQTSGR-----GGLGGQGAGAA-----AAAAAAGG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -------AGQGGYGGLG--SQGTSGRGGLGGQGAGAAA-----AAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 659 CGACCCCTGCAGCGGAGACAGACTGAGCGGCCCGGCACCGCCATG 705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         629 QGA------GAAAAAAAAAAGGAGQGG--YGGLGSQGTSGPG 662
```

NUMBER: 60 1997-10-2 NUMBER: 60 1997-10-2 NUMBER: 60 1997-10-3 NUMBER: 60 1997-10-3 NUMBER: 60 1997-11-1 NUMBER: 60 1997-11-1 NUMBER: 60 1997-11-1 NUMBER: 60	NUMBER: 60 1997-11-2 1997-11-2 1997-12-1 1997-12-1 1997-12-1 1997-12-1 1997-12-1 1997-12-1 1998-03-1 1998-03-1 1998-03-1 1998-03-1 1998-03-1 1998-03-1 1998-03-1	1998-03-31 NUMBER: 66/ 1998-04-01 NUMBER: 60/ 1998-04-01 1998-04-08 NUMBER: 60/ 1998-04-08 NUMBER: 60/ 1998-04-18 NUMBER: 60/ 1998-04-21 NUMBER: 60/ NUMBER: 60/ N
PPLICATION ILING DATE PPLICATION PPLICATION ILING DATE PPLICATION	ATTION THE TOTAL	
PRIOR		

PRIOR FILING DATE: 1996-04-29
PRIOR PILING DATE: 1996-04-29
PRIOR PILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/08414
PRIOR APPLICATION NUMBER: 60/08414
PRIOR PILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084640
PRIOR PILING DATE: 1998-06-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR PILING DATE: 1998-06-07
PRIOR PILING DATE: 1998-06-07
PRIOR PILING DATE: 1998-06-07
PRIOR PILING DATE: 1998-06-15
PRIOR PILING DATE: 1998-06-16
PRIOR PILING DATE: 1998-06-16
PRIOR PILING DATE: 1998-06-16
PRIOR PILING DATE: 1998-06-17
PRIOR PILING DATE: 1998-06-18
PRIOR PILING DATE: 1998-06-07
PRIOR PILING D

		NUMBER: 60/08886: 1998-06-11 NUMBER: 60/08886: 1998-06-11
		ATION NUMBER: 60 DATE: 1998-06-1 ATION NUMBER: 60
		AFLILING DATE: 1998-10-12 APPLICATION NUMBER: 60/089105 FILING DATE: 1998-06-13
	-	NUMBER: 60/08951: 1998-06-16
		NUMBER: 60/08951: 1998-06-16
		NUMBER: 60/08953 : 1998-06-17
	PRIOR FIL	AFFILCATION NOMBER: 00/009398 FILING DATE: 1998-06-17 APPLICATION NUMBER: 60/089653
		FILING DATE: 1998-06-17 APPLICATION NUMBER: 60/089908
~~ <u>~</u>	Query Matcl Best Local Matches 20	Match
οy	1.5	SGRGGLGGQGAGAAAAAAAAAGGAGGGGYGGLGSQGT-SGRGGLGGQGAGAA 66
QQ	152	:
ογ	29	AAAAAAAAGGGGGGGGGGGGGGGGGGGGAGAAAAAAAA
q	212	CATGTGCATGTGTGTGTTGT-GTGCCTGTGTGCACATGTAGAAA-GAAAGTGAT 265
Qγ	127	GTSGRGGLGGQGAGAAAAAAAAAAGGGGGGGGGGGGGGTSGPGGYGPGQQTSGRGGL 184
g	266	GTGTGGGAATGGAGAATTCCAACCCAGGAGACTGTGCCTGGGGCTGCCACGAGAAGT 325
ογ	185	GQGAGAAAAAAAAAAGGAGGGGGGGGGGG
g	326	GGTGACACTTCGAGAAACGGGTTCCAAATGCACGTTGCCTTCCAGTTCTCTGTGTGACCT 385
ογ	217	SGRGGLGG
g	386 T	GGGCAAGTCACTTCAGCTCCCTGAGCTTTGTTTTAAAATATTTTTAAATGTATAAA 445
ογ	234	AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQ 281
g	446	
ογ	282	GGYGGLGSQGT
g	506 6	GTGCAGGAGTGCCTGCTCAGACCTGACACAGGGAACTCTCCCAGAAGGCTGCA 565
ò	320 GG	TOOSTPROCIGSQCTSGPGGYGPGQQT
g	266	GGGCTTCCCTCCCAACCCTCCAACGCCCCACTTGCTGAGCCCCGTGCTTATCTGTTCAGTG 625
ογ	347	GLGGOGGLG 376
윱.	626	
ò	377	SQGTSGRGGLGGQGAGAAAAAAAAAAAGGA 405
g	222 989	CAGCTTGGTGCAAAGCCCCTTGTGTGTGTTTTTTTTTTT
ò	406 G	QGG-YGGLGSQGTAAAAAGG 442
g	746	AAGGTCAGTACTGTCACTTCCCCCACTTTGAAGATGAGGAGAGACACAAATTCTAGATGGA 805
ò	443	AGGGGGGGGGGGGG478
g	908	AGTCACGCAGTGGAAATAGGATCCAGACA

Οy	479	GGAG	482
q	866	TCTTCCATTCCACGCTACGCTTGAAGCTCAATCTCTTCCTTGCTTG	925
οy	483	STOODED A SOURCE STOOD SOURCE	509
q	926	CCCCACCCCCAGATATATCCCATCGCTGCTTGGTGGACAGTAGCCATGACTGGGTTTTGG	985
ογ	510	GGLGGGGAGAAAAAAAAAGGAGGYGGLGSQGTSGRG	548
đ	986	986 TAAAGGIIGCIGAATAAICAGGIIGCIGGTIAGIITITACAITITCACCITICCCAGIGAA	1045
Qγ	549	GLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA	597
qo	1046	1046 ATGGGGCCCCATGAAAAGGCA-GCTCAAGTTGTAAATTACTCAAAGGAAGGAAGAAAG	1104
φ	598		614
đ	1105	1105 GTCTTCTGTTTGCACCTACCCTAAGGATTTGGGGTAGACACTGGGAATTTACTAATTATG	1164
οy	615		655
q	1165	1165 AATTCCAGTGCTTTCCTTGCTGAAAGAGAGGCGTGGAATCAACGCTGAGTGAAGGCATCA 1224	1224
0y	929	QGTSGPGG 663	
q	1225	1225 AGTTTAAG 1232 ·	
Sea: Job	rch com time :	Search completed: December 18, 2002, 16:39:01 Job time : 20.4075 secs	

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
```

- protein search, using sw model OM protein December 18, 2002, 16:13:49 ; Search time 20.1216 Seconds Run on:

(without alignments) 3253.588 Million cell updates/sec

US-09-490-291-4 3493 1 MRGSHHHHHHGSMASGRGGL.....GGYGPGQQTSGIRRPAAKLN Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283224 seqs, 96134422 residues Searched:

Total number of hits satisfying chosen parameters:

length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

hypothetical glyci major ampullate fi fibroin – Chinese hypothetical Description SUMMARIES A36068 T731328 D708070 B70807 A70804 A70804 A70804 A70804 A70801 A70801 B70812 E70895 Length DB % Query Match 1 2135 1386.5 1277.5 1277.5 1228.5 1187 1187 11173 11173 11135 1116.5 1116.5 1016.5 1044.5 1024 1024 1024 1024 1024 108 984.5 984.5 984.5 988.5 988.5 988.5 988.5 Score Š. Result

30 903.5 25.9 694 2 F70868 31 869 24.9 603 2 A70770 33 869 24.6 1588 2 T31611 34 853.5 24.4 465 1 501820 35 833.5 23.4 576 2 A70900 36 817.5 23.4 514 2 G70804 37 804.5 23.1 618 2 A70900 38 804.5 23.1 770 2 S59623 40 800 22.9 501 2 B70531 40 784.5 22.5 747 1 EABO 41 774.5 22.2 1470589 44 774.5 22.2 588 2 F70971 45 76.3 21.8 641 1 QOBE31	hypothetical glyci hypothetical glyci hypothetical glyci		hypothetical glyci hypothetical glyci tropoelastin - she hypothetical glyci	hypothetical glyci hypothetical glyci elastin precursor,	hypothetical glyci hypothetical glyci nuclear antigen EB
903.5 893.5 869.5 869.2 869.2 869.2 87.5 87.5 87.5 800.5	F70868 G70846 A7070	T31611 S01820 A70900 G70804	A70989 S59623 D70931	B70523 H70816 EABO	H70589 F70971 QQBE31
903.5 893.5 869.5 869.2 869.2 869.2 87.5 87.5 87.5 800.5	0000	717	9000	777	777
903.5 893 869 869 860 883.5 883.5 804.5 804.5 776.5 776.5	694 603 603	1585 465 576 584	618 770 639	591 606 747	615 588 641
903.5 893 869 869 860 883.5 883.5 804.5 804.5 776.5 776.5	σφοι	0404	100	منون	0.00
	252	23.44	3333	222	22 22 21 21
01000000000000000000000000000000000000	903.5 893 869	853.5 833.5 817.5	808.5 804.5 804	800 799 784.5	776.5 774.5 763
	30 32 32	3 3 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	33 33 36 36	4 4 4 4 1 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4	4 4 4 4 5 5

ALIGNMENTS

11 1 2 2 2 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	RESULT 1 A36068 major ampullate fibroin protein - orb spider (Nephila clavipes) (fragment) C:Specias: Nephila clavipes C:Accession: A36068 C:Accession: A36068 C:Accession: A36068 A:Title: Structure of a protein superfiber: spider dragline slik. A:Reference number: A36068; MUID:90384959; PMID:2402494 A:Reference number: A36068; MUID:90384959; PMID:2402494 A:Status: preliminary A:Molecule type: DNA A:Residues: 1-718 < CAUA> A:Residues: 1-718 < CAUA> A:Residues: GB:M37137; NID:9159711; PID:9159712 A:Rote: the authors translated the codon GGT for residue 292 as Gln, GTA for residue	Query Match Best Local Similarity 68.3%; Pred. No. 1.1e-112; Matches 495; Conservative 15; Mismatches 39; Indels 176; Gaps 34;	23 OGAGAAAAAAAAAAGGAGGGGGGGGGGGGGGGGGGGGG	GGYGGIGSOGTSGRGGIGGOGAGAAAAAAAAGGAGOGGYGGIGSOGTSGRGGIGGOGA 139 	140 GAAAAAAAAAGGAGGGGGGGGGGGGGGGGGGGGGGGG	200 AGGAGOGGYGGLGSQGTSGRGGLGGOGAGAAAAAAA AGGAGOGGYGGLGSQGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	SQGTSGRGGLGGQGAGAAAAAAAAAGGAGGGGGGGGGGGG	QCTSGRGGLGGGGAGAAAAAAAAGGGGGGGGGGGGGGGGGGGG	GGGGAGAAAAAAAAAAGGAGGGGGGGGGGG	391 AGAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGAGAAAAAA
	pullat S: Nep 08-Mar 10n: A ; Lewi tl. Ac Tl. Ac nce nu 10n: A : prel : prel : prel cs: 1'P	Match ocal S s 495	3 OGAG 1 OGAG	80 GGYG 57 GGYG	0 GAAA G	0 AGGA - - 3 AGGA	25:1 SQGT 1 202 SQG-	290 QGTS 261 QG	349 GGQG 312 GGQG	1 AGAA 1 AG
	RESULT 1 A36068 major am C; Specie C; Date: C; Access R; Xu, M. Proc, M. Proc, M. A; Title: A; Actess A; A	OME	Qy Db	Oy Op	ΟΥ Db	Qy Db	Qy Dp	Oy Dp	oy Pa	Oy Op
D D D D D D D D D D D D D D D D D D D										

	SSAAAAAAAAAA 374 AGAAAAAAAAAG 201 1
130GRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQ	

```
A; Experimental source: strain H37Rv
                                                                                                                                                     g
                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                            à
                                                                                                             셤
                                                                                                                                 à
                                                                                                                                                                         à
                                                                                                                                                                                           a
                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                  1006 DGASGLGLGLSGFDGGQGGQGGAGGSAGAGGINGAGGAGGNGGDGGDGATGAAGLGDNGG 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1066 VGGDGGAGGAAGNGGNAGVGLTAKAGDGGAAGNGGNGGAGGAGGAGGAGDNNFNGGQGGAGGQ 1125
                                                                                                                                                                                                                                                                                                                                                                                                         949 GTGGKGGAGGAGTLGVGGSG---GTGGDGGDAGSGGGGFGGAAGKAGGGGGNGGRGGDGG 1005
                                                             166
                                                                                                                                                                                                      708
                                                                                                                                                                                                                          360
                                                                                                                                                                                                                                     541
                                        471
                                                                                531
                                                                                                                      532 FSGGAGGAGGAGGNTGVGGTNGSGGQG-GTGGAGGAGGAGGVGADNPTGIGGTGGTGGKG 590
                                                                                                                                                                                                                                                                   391
                                                                                                                                                                                                                                                                                                                    481
                                                                                                                                                                                 270 AAAAAAAAGAGGAGGYGG-----LGSQGTSGRGGLGGQ-GAGAAAAAAAAAAAAGGA---- 317
                                                                                                                                                    472 TGGNAGAGGAGGAAGTGGTGGVVGAAGKAGIGGTGGOGGAGGAGGAGSAGTDATATGATGGTG
                                                                                                                                                                                                                          GQGGYGGLGSQ----GTSGPGGYGPGQQTSGRGGLGGQGA------GAAAAAA
                                                                                                                                                                                                                                                                              GQGGYGGLGSQGTSGPGQYTSGRGGLGGQGAGAAAAAAAAAAAGGGGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      588 LGGQ-GAGAAA-----AAAAAAAGGAGQGGYGGLGSQG----TSGRGGLGGQ
GKGGVGGVAGLGGAGGAAGQLFSAGGAAGAVGVGGTGGQGGAGGAGAAGADAPASTGLTG
                                       GTGFAGGAGGVGGGGGAIAGGINGSGGAGGTGGQGGAGGAGGAGGAGGAGG
                                                                                                  ----SGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQ-----GGYGGLGSQG
                                                                                                                                           TSGRGGLGGQG--AGAAAAAAAAAGG-AGQGGYGGLGSQGT---SGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                         GAAAAAAAAAAAAGGA-GQGGYGGLGSQGTSG---RGGLGGQ-GAGAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                               AAAAAAAGGAGQGGYGGLGSQG----TSGRGGLGGQGA----GAAAAAAAAAAAAGGA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GA----GAAAAAAAAAAGGA----GQGGYGGLGSQGTSGPGGYG 665
                                                                                                                                                                                                                                                                361 AAAAGGA-GQGGYGGL-----GSQGTSGRGGLGGQGA---
                                                            QG--AGAAAAAAAAGG-
                                                                                                                                                                                                      649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              630
                                      112
                                                                                                                                                                                                                                            402
                                                                                                                                                                                                                                                                                     169
                                                                                                                                                                                                                                                                                                         392
                                                                                                                                                                                                                         318
                                                                                                  167
                                                                                                                                                                                                                                                                                                                                                434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1126
                                                                                                                                                            g
                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
셤
                                     g
                                                          ò
                                                                             g
                                                                                                 δ
                                                                                                                  g
                                                                                                                                         ò
                                                                                                                                                                               ò
                                                                                                                                                                                                   g
                                                                                                                                                                                                                         òγ
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                     ογ
                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
Wypothetical glycine-rich protein Rv3514 - Mycobacterium tuberculosis (strain H37RV) C; Species: Mycobacterium tuberculosis C;Date: 17-Ju1-1998 #sequence_revision 17-Ju1-1998 #text_change 20-Jun-2000 C; Accession: D70807 C; A; Feltwell, T.; Gentles, S.; Hamin, N.; Holroyd, S.; Connor, R.; Davies, R.; Devilin, K.; Feltwell, T.; Gentles, S.; Hamin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nathors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A; Title: Deciphering the blology of Mycobacterium tuberculosis from the complete genome A; Reference number: A70500; MUID:98295987; PMID:9634230 A; Reterence number: A70500; MUID:98295987; PMID:9634230 A; Reterence number: A70500; Multiple Reduction of Shown A; Molecule type: DNA A; Mesidues: 1.1489 CCOL> A; Residues: 1.1489 CCOL> A; Cross-references: GB:AL022022; GB:AL123456; NID:93261554; PIDN:CAA17751.1; PID:9292445
```

RESULT 5 B70807 hypothetical glycine-rich protein Rv3512 - Mycobacterium tuberculosis (strain H37RV)

```
36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         970 TGGTGGKGGTGGAGDDSAGGTGGTGGAGGNAGAGGLANTGGTAGNAGIGGDGGQGGNGGQ 1029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           157
                                                                                                                                                                                                                                                                                                                                               435
                                                                                                                                                                                                                                                                                                                                                                                                                       436 SGGAGGOGGAGGAGADNPTGIGGTGGDGGTGGAAGAGGAAGTGGTGGMIGTTGN 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGAGGQGGDGGAGGAGADADQPGATGGTGFAGGAGKAGGSSSAGGTNSSGSAGGTGR 675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  386 LGGQGAGAAAAAAAAAAAGGGGYGGLGSQGT---SGRGGLGGQGAGAAAAAAAAAAAGG 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SSCAGGTNGSGGAGGTCGQVVAGGAGISFSNGSNGGTGGTGGVG--GTGGDGGNAGTGAG 912
                                                                                                                                                                      AAAAAAAAGA-GQGGYGGLGSQ-----GTSGRGGLGG----QGAGAAAAAAG 113
                                                                                                                                             16 GRGGLGGQG-----AGAAAAAAAAGGAGQ-----GGYGGLGSQGTSGRGGLGGQGAGA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YGGLGSQCTSGPGGYGPGQQTSGRGGLGGGGAQAAAAAAAAAGAGGGYGGYGG-LGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AGQGGYGG-LGSQGTSGRGGLGGQ----GAGAAAAA------AAAAGGA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQGGYGGLGSQGTSGPGG-YGPGQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TSGRGGLGGQGAGAAAAAAAAAAAGGAG-QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                        378 AAGNAGVGGAGGQG--GDGGAGGAGADADQPGATGGTGFAGGAGGAGGAGGSSGAGGTNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                SGRGGLGGQ-----GAGAAA-----AAAAAAGGA------GQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     496 AGVGCAGGGGGGGGGGGAGGADADQPGATGGTGFAGGAGGAGGAGGSSGAGGTNGSGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LGSQG------TSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGG-LGSQGTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGLGGQ----GAGAAAAA-----AAAAGGAGQ------GGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q-GTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGYGG-LGSQGTSGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    565 AAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGA-----GQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTSGRGGLGGQ------GAGAAAAA-----AAAAAGGAGQGGYGGLG-----S
                                                                                                        Indels 211;
                                                                 Length 1489;
                                                                 DB 2;
                                                                                                      36; Mismatches 267;
                                                                                        . 2e-63
                                                               35.9%; Score 1254.5; 39.5%; Pred. No. 2.2e
                          Superfamily: collagen alpha 1(IV) chain
                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGTGGTGGVG 1098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 656 QGTSGPGGYG 665
                                                                                    Similarity
                                                                                                      Matches 336;
A; Gene: Rv3514
C; Superfamily:
                                                                 Query Match
                                                                                      Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  619
                                                                                                                                                                                                                         99
                                                                                                                                                                                                                                                                318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

H37RV)

```
140
                             g
                                                               ò
                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
C; Species: Mycobacterium tuberculosis
C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C; Accession: B70807
R; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Haminh, N.; Holroyd, S. Rajanfream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A; Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A; Reference number: A70500; MUID:98295987; PMID:9634230
A; Status: preliminary; nucleic acid sequence not shown; translation not shown A; Molecule type: DNA
A; Cooss-references: GB:AL022022; GB:AL123456; NID:93261554; PIDN:CAA17749.1; PID:9292444
A; Exceptimental source: strain H37Rv
C; Genetics:
A; Genetics:
A; Genetics:
C; Superfamily: collagen alpha 1(1) chain; fibrillar collagen carboxyl-terminal homology;
                                                                                                                                                                                                                                                                                                                                                                                                           29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    275 TGGTGGNGGNGAIGGAGGDAGGSGNSGNGGIGGKGGNAGAGGAAGSNGGTVGANGTGGD 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        226 G--AGAAAAAAAAAAAAAAAAA 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            273 AAAAGGAG------QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGL 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GSQCTSGP---GGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSG 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            430 GAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGYGGL 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   571 AGGTGGDRTVGGGTVPAGSGGQG--GNAGGGGAGGQ--GGADGGSGGDGGDAGTGGNGGN 626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          627 GGNRNSGNGTGGAGGNGGCGANGGAGGAGGSGGGTGGNGGAGGDAGDAGNGCNGNCTCNG 686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               687 GNGGNGGIAGMGGNGGAGTGSGNGGNGGSGGNGGNAGMSGNSGTGSGDGGAGGNGGAAGT 746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGL----GSQGT-SGRGGLGGQGAGAAA 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  66
                                                                                                                                                                                                                                                                                                                                                                                                                                           14 ASGRGGLGGQGAGAAAAAAAAAAGACQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 215 PGANGIAGNGGDGGDGAAGAVGISGATGAGDGGHGGTGAAGGNGGTGGAGGSGIDGVGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----AGQGGYGGLG-----SQGTSGRGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100 GA-GAAAAAAAAAGGA----GOGGYGGLG-----SQGTSGRGGLGGQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------GAGAAAAAAAAAAAGAGGGYGGLGSQ----GTSGPGGYGPGQQTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RGGLGGQGAGAAAAAAAAGGAG-QGGYGGLGSQ------GTSGRGGLGGO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             G----SQGTSGPGGYGPGQQTSGRGGLGGQGAGAAA-AAAAAAAGGAGQGGYG-GLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAAAAAGGAGQGGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     393 VGGAGGAGGSGPNTSPGGNGGQGGGGGGGGGG-GAAGAGGAGGANGTAGNGGQGGAGGT
                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 40.09
Matches 318; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              74 AGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
```

```
Cypecies: Mycobacterium tuberculosis
Cyaccesion: A70934
Rydole, S.T.: Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamiln, N.; Holroyd,
Rajandram, M.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A; Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A; Rateronce number: A70500; MUID: 98295987; PMID: 9634230
A; Reference number: A70500; MUID: 98295987; PMID: 9634230
A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Residues: 1-1306 <CCL>
A; CCCS-references: GB: AL021942; GB: AL123456; NID: 93242298; PIDN: CAA17449.1; PID: 9290
A; Genetics:
A; Genetics:
A; Genetics:
A; Genetics:
A; Genetics:
A; Genetics:
A; Cyperfamily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                hypothetical glycine-rich protein Rv0578c - Mycobacterium tuberculosis (strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               747 GGTGGDGGLTGTGGGGGGG-GTGGDGGNGADNTANMTAQAGGDGGNGGDGGFGGG 804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           395 AAAAAAAAGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGAGGGGYGG---- 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 GSHHHHHHG-----SMASGRGGLGGQGAGAAAAAAAAAGG----AGQGGYGGLGSQGTS 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          506 NASAPGAAGSPGGNGGNGGPGGAGGEGGAGGLALAASGQNGSQGAGGDGGAGGNGGTPGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---GAAAA-----AAAAAAGGAGGGYGGLGSQGTSG--PGGYG-PGQ--QTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            265 AGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAG-GAGQGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        353 -AGAAAAAAAAAAGGA-------GQGGYGGLGSQGTSGRGGLGGQGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54 GRGGLGGQGAGAAAAAAAAAGG-AGQGGYGG--LGSQGTSGRGGLGGQGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                102 GAAAAAAAAAAGG-----AGQGGYGGL------GSQGTSGRGGLGGQGA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   187 QGAGAAAAAAAAAAGG------AGQGGYGGLGSQ------GTSGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  222 LGGQGAGAAAAAAAAAGG-AGQGGYGGLGSQ-------GTSGRGGLGGQG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       804 GKGGRGGTGGPGGAGPAGTTGSQGAGGNGGSGGTGGDPGDGGNGANGSVFTNNGIGGNGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 194;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     324 GLGSQ-GTSGPGGYGP----GQQTSGRGGLGGQG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34.3%; Score 1197; DB 2;
38.1%; Pred. No. 3.2e-60;
tive 38; Mismatches 296;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 38.1% Matches 325; Conservative
                                                                                             658 ISGPGGYGPGQQTSG 672
                                                                                                                                                              805 GAGGGGLTAGANGTG 819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
```

q	924 NGSLRGSGGAGGHGGNGGNASRGMGGDGCTGGAGGNAGQIGNGGAGGNGGDGCTGSDGNP 983	
ò		Oy 288 GSOGTSGRGGLGGQGAGAAAAAAAAAAAAGGGGGGGGGGGTS 331
g	: IGATGATGTFDAG	Db 519 GGNGGNGTGGVNGADNTLNPDTPGGAGEPCGAGGAGGAGGAAGGPGGTGGTGGNGGNG 576
ογ	497 PGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGGYGGLGSQGTSG 546	332 GPGGYGPGQQTSGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGSQGTS :
qq	1042 ADGHGGNGGTGGVGGTGGAGGGGGGGGGGGKALSPTGNNGSQGAGGDGGAGGAGG-GTGG 1099	5// GNGGNCGNGGNGGNAGNNSTNAPVGGEG-CAGGDGCAGGAGGANGCTAGSQGTG
Š Š	547 RGGLGGGGA-GAAAAAAAAAAGGAGGGGGGGGGGGGGGGG	Qy 382 GRGGLGGQGAGAAAAAAAAGAGGGGGGGGGGGGGGGGGGG
a a	TIOU TGGDGGKGANGTEENSSLAGTGGTGGTGGTGGTGGTGGGGGGGGGGTGSTLGATGANGAGTTGTTIN	Qy 438 AAAGGAGOGGYGGLGSOGTSGRGGLGGOGAGAAAAAAA 479
S 8	PAGNGGVGGSGGLGSAFGPGGTGGMG-GAGGTSTVSAGGDGGRGGFGGDGLDASSGGNGG	Db 684 GGTGGSGGNGGDGGAGGIGGAGGNGIPGTGTEPAGGTGAKGGDGGDGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
οy		
qq	1217 DGGHGGDGFRTAG 1229	/44 GQGGGAGGAGGAAYIKGDGYGAAYAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGG
RESULT	7 LT 7	DD 801 IGGGAGGTGGSGGHAGKGGAGGTTTTVPGNGGNAGGNGGGNAGAGGGGG 857
hype C; Sr	hypothetical glycine-rich protein Rv3507 - Mycobacterium tuberculosis (strain H37RV) C;Species: Mycobacterium tuberculosis C;Species: 17-11-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000	QY 573 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGGG-QGGYGGLGSQGTSGRGGLGGQGA 631
	cession: £/0800 6. S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S nnor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.	Oy 632 -GAAAAAAAAAAGGAGGYGGLGSQ-GTSGPGGYGPGOOTSG 672
Raja		Db 915 HGTVGAQFVPATSLPTPNCGAGGNGGTGSNGGAPGPAG-APGPTTGG 960
A; At A; Ti A; Re	A.Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A.Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A;Reference number: A70500; MUID:98295987; PMID:9634230	RESULT 8
A; Ac A; St A; Mc	A;Accession: E70806 A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA	hypothetical glycine-rich protein Rv2490c - Mycobacterium tuberculosis (strain H37RV) C; Species: Mycobacterium tuberculosis C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text change 20-Jun-2000
A; Re A; Cr	A;Residues: 1-1381 <col/> A;Cross-references: GB:AL022022; GB:AL123456; NID:93261554; PIDN:CAA17744.1; PID:9292444	C; Accession: A70869 R; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
A; Ge		S. HO
C; St	C;Superfamily: collagen alpha 1(IV) chain	A.Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A.Title: Deciphering the hiology of Myccharterium tuberculosis from the complete geno
ŏ#:	33.9%; Score 1182.5; DB 2; Length 1381; illarity 38.8%; Pred. No. 2.2e-59;	A; Reference number: A70500; MUID: 98295987; PMID: 9634230 A; Accession: A70869
Ä	tches 321; Conservative 32; Mismatches 275; Indels 199; Gaps 37;	A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA
9 P	16 GRGGLGGGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQ 61	A; Residues: 1-1660 <col/> A; Cross-references: GB:AL021246; GB:AL123456; NID:93261507; PIDN:CAA16067.1; PID:9279 A; Experimental source: strain H37Rv
δ d	62 GAGAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	C;Genetics: A;Gene: Rv2490c C;Superfamily: collagen alpha 1(IV) chain
3 8	QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGSQGTSGP	5, 8,4e-59; Length 1660;
g	281 NGGAGGLFMNGGDGGAGGQGGDGAAGDAAASAGGTGGKGGQGGDGGTGGAGGPVLF 338	SSI; CONSELVA
6 G	170GGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGY 208	1. GSFASSONGGEGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
δŏ	GGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGT	OY 47 -LGSQGTSGRGGLGGQ-GAGAAAAAAAAGGGGGGCYGGLGSQ 88
셤	GGVGAGGTGGTGGTGGNGADAAAVVGFGANGDPGFAGGKGGNGGIGGAAVTGGVAGDGGT	Oy 89 GTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGL 123
oy Oy	255 SERGELGGG	Db 571 GGNGGDGGAGGRGGDGGAG-GAGGDAPAGRAGSQGVGGDGGAGGAGGAPGNGGSGGRGDM 629

A;Gene: Rv3345c C;Superfamily: collagen alpha 1(IV) chain	Query Match 33.1%; Score 1157; DB 2; Length 1538; Best Local Similarity 34.4%; Pred. No. 6.2e-58; Matches 317; Conservative 36; Mismatches 308; Indels 260; Gaps 31	QY 1 MRGSHHHHHGSMASGRGGLGGQCAGAAAAAAAAGGAGQGGYGGLGSQGTS 53 :	Qy 54 GGGLGGQGA-GAAAAAAAAAGGGGGGGGGGGGGGG	92 GRGGLGGGGAGAA : : 669 GNGGLGGKGGDAF	Qy 140T 178 GAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	QY 179 SGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 236	Qy 237 AAGAGQGGGGGGSQGTSGRGGLGQQ-CAGAAAAAAAAGGACQGGGGGG 286	Qy 287 16SQGTSGRGGLGGQ	QY 302GAGAAAAAAAAAGGAGOGGYGGLGSQGTS 331	OY 332 GPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAGAGGGGGGGGGGGGGGGGGGG	Qy 388 GQGAGAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAA 436	Qy 437 AAAAGGAGQGGYGGLGSQGTSCRGGLGGQGAQAAAAAAAAAAGGG 484	Qy 485 GYGGLGSQGTSGPGGYCGGGGGGGAAAAAAAAGGAG 532 	Qy 533 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTS-GRGGL 588	QY 589 GGGGAGAAAAAAAAAAGGAGGGGGGGGGGGGGG	OY 636GAAAAAAGGAGGGGYGGLGSQGTSG 660 Db 1375 DGGNGGNGGNSGNSTGVAGLAGGAAGAGGNGGGTSSAAGHGGSGGSGGSGGTTGGAGAAGG 1434	Qy 661 PGGYGPGQQTSGIRR 675 	RESULT 10
Oy 124	148 AAAGGAGGGYGGL-GSQGTSGPGYG		OY 243 QGGYGGLGSQCTSGRGGLGGQGAAAAAAAAGGA-GQGGYGGLGSQGTSGRGGLGGQ 301	Oy 302 GAGAAAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGG	Qy 348 LGG-OGAGAAAAAAAAAACGAGQGGYGGLGSQ		OGGY 11 DGGV		GSQG : GNAG	LGGLT	AAGG - DAGM		RESULT 9	hypothetical glycine-rich protein Rv3345c - Mycobacterium tuberculosis (strain H37RV) C; Species: Mycobacterium tuberculosis C; Date: 17-7u1-1998 #sequence_revision 17-7u1-1998 #text_change 20-Jun-2000 C; Accession: H70846	R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544.	A; Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A; Reference number: A70500; MUID:98295987; PMID:9634230	A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA A;Molecule type:	A/Cross-references: GB:ALDZ1841; GB:ALJZ3456; NID:g3261517; PIDN:CAA17117.1; PID:g326151 A/Experimental source: strain H37Rv C;Genetics:

QY 54.3 GTSGRGGLGGQGAGAAAAA	RESULT 11 A70896 hypothetical glycine-rich protein Rv1091 - Mycobacterium tuberculosis (st C; Species: Mycobacterium tuberculosis C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 24-Nov-15 C; Accession: A70896 C; Accession: A70896 R; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harrie; Connor, R.; Davies, R.; Devilin, K.; Feltwell, T.; Gentles, S.; Hanlin, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Square Nature 39, 357-544, 1998	A;Title: Deciphering the biology of Mycobacterium tuberculosis from the carbive fraction and the carbive fraction of the carbive fraction and the	Ouery Match 32.6%; Score 1139; DB 2; Length 853; Best Local Similarity 40.3%; Pred. No. 4.1e-57; Matches 298; Conservative 27; Mismatches 304; Indels 110; Gaps Oy 11 GSMASGRGGLGGOGAGAAAAAAAAGGAGGGGGGGGGGGGGGGGG	QY 108 AAAAAGGACQGGYGGLGSQGTSGRGGLGGGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	GKGGSGLSGNANGGAGGD GLGGGGAAAAAAA GDGGAAAGTGIAGAGGRG AAAAAAAGGA AGGAGGGAGGGGCNGCNG AAGGAGGGGGCGCNGCNG AGGAGGGGGGGGGCGCNGCNG AGGAGGGGGGGGGG
hypothetical glycine-rich protein Rv1450c - Mycobacterium tuberculosis (strain H37RV) C; Species: Mycobacterium tuberculosis (Species: Mycobacterium tuberculosis (Strain H37RV) C; Species: Mycobacterium tuberculosis (Species: Mycobacterium tuberculosis from the complete genome My Mytitle: Deciphering the biology of Mycobacterium tuberculosis from the complete genome MyReference number: MyOSOO; MuID:98295987; PMID:9614230	A; Accession: E70917 A; Status: preliminary; nucleic acid sequence not shown; translation not shown A; Status: preliminary; nucleic acid sequence not shown; translation not shown A; Rosidues: 1-1329 <col/> A; Rosidues: 1-1329 <col/> A; Cross-references: GB:295844; GB:AL123456; NID:g3250713; PIDN:CAB09271.1; PID:g2131046 A; Experimental source: strain H37Rv C; Genetics: A; Gene: RV1450c C; Superfamily: collagen alpha 2(1) chain; fibrillar collagen carboxyl-terminal homology Query Match Best Local Similarity 35.8%; Score 1145; DB 2; Length 1329; Best Local Similarity 35.9%; Mismatches 300; Indels 242: Gaps 35; Matches 323; Conservative 35; Mismatches 300; Indels 242: Gaps		151 GGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	Db 436 GNGAPGAPFSGGDPNGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	

```
Soch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon les, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. 4, 1998

V. R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. ing the biology of Mycobacterium tuberculosis from the complete geno sec. A70500; MUID:96295987; PMID:9634230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : GB.AL021897; GB:AL123456; NID:93256022; PIDN:CAA17207.1; PID:e125.rce: strain H37Rv
                                                                                                                                                                                                                                                                                                                                 ine-rich protein Rv1091 - Mycobacterium tuberculosis (strain H37RV)
terium tuberculosis
38 #sequence_revision 17-Jul-1998 #text_change 24-Nov-1999
16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ary; nucleic acid sequence not shown; translation not shown
                                                                                                                                                                                                                     AAAAAAAAAGGAGQ----GGYGGLG------SQGTSGRGGLGGQ 629
                                                                                                                                                                                       -AGAAAAAAAAAAGGAGGGYGG-LGSQGTSGPGGYGPGQQTSGIRRPAAKL 680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPGVAGGNGGAGGSAGLWGTGGAGGDGGNGRSGPVNVAGSAGGNGGAGGAG 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GQGGYGGLG-----SQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYG 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAAAAAAAGGA-----GQGGYGGLGSQGTSG----RGGLGGQGAGAAAA 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGQGAGAAAAAAA-------AAAGGAGQGGYGGLGSQCTSGRGG 587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      32.6%; Score 1139; DB 2; Length 853;
40.3%; Pred. No. 4.1e-57;
Live 27; Mismatches 304; Indels 110; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLGGQGAGAAAA----AAAAAAGGAGQGGYGGLGSQGTS-----GRGGL, 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GNGGKGGAGGAAFSINFTAGDGGAGGAGGSGGHALLWGAGGAGGNGGSGGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TSGPGGYGPGQQTS----GRGGLGGQGA----GAAAAAAAAAAGGAGQGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GRGGLGGQG-----AGAAAAAAAAAAGGAGQGGYGGL-----GSQGTSGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGG-----AGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ssigned collagens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 arity 40.3%
```

588 500

```
Ricole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon S.Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A.Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A.Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A.Accession: A70812
A.Stefarence number: A70500; MUID:98295987; PMID:9634230
A.Stefarence number: A70500; MUID:98295987; PMID:9634230
A.Stefarence number: A70615
A.Stefarence number: A70615
A.Stefarence number: A70615
A.Stefarence number: A70616
A.Stefarence number: A70616
A.Stefarence number: A70616
A.Stefarence number: A70617
A.Stefarence n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    hypothetical glycine-rich protein Rv0833 - Mycobacterium tuberculosis (strain H37RV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C; Species: Mycobacterium tuberculosis
C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
                                                                                                                                                                                                                                                                                                                                   501 AGYGPGSAVAASA----GAGSAGYGP-GSQASAAASRLASPDSGARVASAVSNLVSSGPT 555
         273 G------SAAAAAAAAAAGPGGYG----PGQQGPGGYGPGQQ------GPSGAGS 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            122 GGTGGGAGSGGTGGVG--GTGGAGGEFASGAGGSGGTGGAGGTGGAGGLFCAGGAG 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             238 AGGSGGAAGTFGDTGNSGGAGGAGGKAGILFGSGGAGGSGGAGGFANGSTGGAGGAGGA 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            56 GGLGGQGAGAAAAAAAAAAGGA-----GQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   589 GGQGAGAAAAAAAAAAGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              529 GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --AGPGGYGP-AQQGPSGPG-----PAQQGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        180 GLG--GQGNHTGGHGGAGGSAGLLALGDGGAGGAGGAATTGTGGAGGAGGKAGLLFGSGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  311 AAAAAAAGPGQGGLGGYGP-GQQGPGGYGPGGYGPGSASAAAAAAAGPGQQGPGGY
                                                                                                                                                                                 411 GGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSG--RGGLGGQGAG
                                                                                                                                                                                                                                                                                                     469 AAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAA
                                                                GPGCYGPGQQTSGRGGLGGQ-----GAGAAAAAAAAAAAGGA-GQGGYGGL--GSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 135;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31.9%; Score 1114.5; DB 2; Length 749; Ilarity 39.6%; Pred. No. 8.7e-56; Conservative 18; Mismatches 309; Indels 135;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16 GRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQ-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           649 GYGGLGSQ-----GISGPG 662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     556 SSAALSSVISNAVSQIGASNPG 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 303; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Gene: Rv0833
C;Superfamily: elastin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C; Accession: A70812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 168
                                                                   à
                                                                                                                       셤
                                                                                                                                                                                 ò
                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    spidroin 2, dragline silk fibroin - orb spider (Nephila clavipes) (fragment)
N.Alternate names: silk fibroin, dragline
C.Specias: Nephila clavipes
C.Specias: Nephila clavipes
C.Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Mar-1998
C.Accession: A44112; S27824
S.Hinman, M.B.: Lewis, R.N.
J. Biol. Chem. 267, 19320-19324, 1992
A;Title: Isolation of a clone encoding a second dragline silk fibroin. Nephila clavipes
A;Reference number: A44112; MUID:92406876; PMID:1527052
A;Status: preliminary
A;Accession: A44112
A;Retas: preliminary
A;Molecule type: mRNA
A;Residues: 1-627 <HINN
A;Cross-references: GB:M92913; NID:9159713; PID:9159714
A;Note: sequence extracted from NCBI backbone (NCBIP:113893)
B;Hinman, M.B.: Lewis, R.N.
submitted to the EMBL Data Library, May 1992
A;Description: Isolation of a clone encoding a second dragline silk fibroin: Nephila cla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         36;
                                                                                                                          534
                                                                                                                                                        654
                                                                                                                                                                                                                                                                                                                                                                                          769 GGVLYGNGGAG--GQGSSGGIGGPGAT--GGAGGKGGDGGDAQLIGDGGNGGAGGTG 824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGLGGQGAGAAAAAAAAGGAGQGGYGCLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGA 355
GLGGQGA--GAAAAAAAAAAGGAG-0GGYGGLG----S0GTSGRGGLGGO-GAGAAAAA 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  186 -GQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRG----GLGGQGAGAAAAAAAAAGG 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     218 QGPGGYG----PGQQGPGGYGPGQYGLSGPGSAAAAAAAAAAAAGPGQQGPGGYGP----PGQQGPSGP 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              77 AGQGGYGGLG-SQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQ----GTSGR 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLG----- 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241 AGGGGYGGLGSQGTSGRGGLG-GQ----GAGAAAAAAAAAGGAGGGGYGGLGSQGTSGR 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          535 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                  595 AAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLG
                                                                                                                475 AAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQG
                                                          604 GAGGNAGLLGVGGAGGHGASGGAGDRGGAGGTGFISSDGGAGGDGGDGGNGGAGGTGGLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32; Mismatches 180; Indels 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         32.5%; Score 1135; DB 2; 47.5%; Pred. No. 5.5e-57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         655 SQGTSGPGGYGPGQQTSGI 673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                825 -- GTPGPGGPGGSGGLGGL 841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Cross-references: EMBL: M92913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 47.5 Matches 324; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: S27824
A; Molecule type: mRNA
A; Residues: 19-627 <HI2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
   423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        296
                                                          合
                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
   à
                                                                                                                    ò
                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
```

32;

								_				
70 196	122 255	178 314	226 371	279 431	325 490	380	434	487 658	532	592 774	646 834	
. GSWASGRGGLGGQCAGAAAAAAAAGGAGOGGYGGLGSGGTSGRGGLGGGGGAGAAAAA	. AAAAGGAGGGGGGGGGGGGGGGGGGGGGGGGGAAAAAA	LGSQGTSGRGGLGGQGAGAAAAAAAAAGGGGGGGGGGGGGGGGGGGG	SGRGGLGGGGGAGAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGGG 	AGAAAAAAAAAAAAGGGGGGGGGGGGGGGGGGGGGAAAAA	GQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGL 	GSQGTSGPGGYGPGQQTSGRGCLGGQGAAAAAAAAAAAAGGGGGGGGGGGGGGTGT	. SGRCGLGGOGGGGAAAAAAAAAGGGGGGGGGGGGGGGGG	AAAAAAGGAGGGGGGGGGGGGGGGGGAGAAAAAAAAAGGGG		OGGYGGLGSOGTSGRGGLGGGGAQAAAAAAAAGGGGGGGGGGGGGGGGGGGGG	AGAAAAAAAAAAGGAGGGGGGGGGGGGGGGGGGG	QGGYGGLGSQGTSGPGGYGPG 667 DTTIVVNGMLAPSTLGYGGNGVNGGAG-GTG 867
11 139	71 197	123 256	179 315	372	280	326	381 548	435	488 659	533 718	593 775	647 835
Oy Dp	ç q	QY Db	Oy Dp	oy Dp	Oy Dp	Qy Db	o o	Qy Db	oy Ob	Qy Db	Oy Db	oy do

Search completed: December 18, 2002, 16:22:06 Job time : 29.1216 secs

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
```

OM protein - protein search, using sw model

December 18, 2002, 16:06:04; Search time 10.4528 Seconds (without alignments) 2702.183 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-490-291-4 3493 1 MRGSHHHHHHGSMASGRGGL......GGYGPGQQTSGIRRPAAKLN 681

Scoring table:

BLOSUM62 Gapop 10.0 , Gapext 0.5

112892 Total number of hits satisfying chosen parameters:

112892 seqs, 41476328 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_40:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARTES

		d			SUMMARIES		
Result		Query					
No.	Score	Match	Match Length	BB	qi	Description	ption
1	2135	61.1	747	٦	SPD1_NEPCL	P19837	
7	1277.5	36.6	1901	7	YZ08_MYCTU	053553	mycobacteri
e	1168	33.4	5263	٦	FBOH_BOMMO	P05790	bombyx mori
4	1135	32.5	627	٦	SPD2_NEPCL	P46804	nephila cla
S	1106.5	31.7	914	7	WA22_MYCTU	006794	mycobacteri
9	1044.5	29.9	778	٦	YQ34_MYCTU	P71933	
7	1028	29.4	957	Н	X278_MYCTU	P56877	_
8	988	28.3	801	٦	Y747_MYCTU	053810	_
6	916	27.9	860	Т	ELS_MOUSE	P54320	_
10	626	26.9	864	П	ELS_RAT	099372	rattus norv
11	869	24.9	603	П	YD25_MYCTU	010637	mycobacteri
12	853.5	24.4	465	7	GRP2_PHAVU	P10496	
13	775.5	22.2	747	Н	ELS_BOVIN	P04985	bos taurus
14	763	21.8	641	-	EBN1_EBV	P03211	epstein-bar
15	748.5	21.4	1356	7	CA21_ONCMY	093484	oncorhynchu
16	704.5	20.2	750	П	ELS_CHICK	P07916	
17	698.5	20.0	1355	Н	CA21_RANCA	042350	
18	691	19.8	730	٦	ELS_HUMAN	P15502	
19	989	19.6	498	П	Y118_MYCTU	050615	
20	619	19.4	1733	7	VNUA_PRVKA	P33485	
21	678	19.4	543	7	YP91_MYCTU	020630	mycobacteri
22	673	19.3	384	П	GRP1_PETHY	P09789	petunia hyb
23	699	19.2	481	-	LORI_MOUSE	P18165	mus musculu
24	655	18.8	1027	7	CAFF_RIFPA	P30754	riftia pach
25	644	18.4	338	Н	GRP_ARATH	P27483	arabidopsis
56	626.5	17.9	1262	٦	CA13_CHICK	P12105	gallus gall
27	626	٠	491	П	YK98_MYCTU	010707	mycobacteri
28	626	٠	515	٦	YI40_MYCTU	050594	mycobacteri
29	625	17.9	1372	7	CA21_MOUSE	001149	mus musculu
30	621	٠	1466	7	CA13_HUMAN	P02461	homo sapien
31	608	17.4	1372	7	CA21_RAT	P02466	rattus norv
32	608	17.4	1453	7	CA11_CHICK	P02457	gallus gall
33	607.5	17.4	1418	П	CA12_HUMAN	P02458	homo sapien

09xsj7 canis famil P04258 bos taurus 053416 mycobacteri P04525 homo saplen P08123 homo saplen P28481 mus musculu P08121 mus musculu P046592 canis famil P02465 bos taurus P11087 mus musculu P04647 gailus gall P17140 caenorhabdi
CA11_CANFA CA13_BOVIN YA68_MYCTU CA11_HUMAN CA12_HUMAN CA12_MOUSE CA12_CANFA CA21_CANFA
анненнанна
1466 10460 4633 1464 11464 11366 11366 11364 1758
17.3 17.3 17.2 17.1 17.1 17.1 17.0 16.0 16.4
604.5 603.5 602.5 598 598 597.5 597 597 595 595
######################################

ALIGNMENTS

RESULT 1 DAT DIPER DAT O1-198 DAT
--

```
RV3508
                                                                                                                                                                                                                                                                                                                                                             1901 AA; 147627 MW;
                                                                                     (Rel. 39, Created)
                                                                                                                                                                                                                                                                                                       EMBL; AL022022; CAA17745.1; -.
                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 55., Matches 351; Conservative
                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                          30
                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                                                                                                 SEQUENCE FROM N.A.
                               655
                                           641
                                                                                                                                                                                                                                            SUBFAMILY.
                                                                                    30-MAY-2000 (
30-MAY-2000 (
16-OCT-2001 (
                               GGLGS
                                           637 GGVGS
                                                                         YZ08_MYCTU
053553;
                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                         Query Match
                               651
       591
                                                                                                                                                                                                                                                                                                                                                                                                             292
                                                                                                                                                                                                                                                                                                                                                                                                                                      352
                                                                                                                                                                                                                                                                                                                                                                                                                                                  90
                  셤
                               à
                                          qq
                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                        34;
                                                                                                                 V -> L (IN REF. 1).
S -> T (IN REF. 1).
NPGLSGCDVLIQALLEVVSALIQILGSSSIGQVNYGSAGQA
                                                                                                                                                                                                                                                                                                260
                                                                                                                                                                                                                                                                                                                                    390
                                                                                                                                                                                                                                                                                                                                                            450
                                                                                                                                                                                                                                                                                                                                                                        419
                                                                                                                                     IQIVGQSVYQAL -> ILVFLDVMSSFKLFSRLFLLLSRS
                                                                                                                                                                                                           80 GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 139
                                                                                                                                                                                                                  348
                                                                                                                                                                                                                                                                                                                                                                                           550
                                                                                                                                                                                                                                    AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA.-----AGGAGQGGYGGLG 250
                                                                                                                                                                                                                                                                        201
                                                                                                                                                                                                                                                                                   SQCTSGRGLGGQGAGAAAAAA----------AAAAGGAGQGGYGGLGS 289
                                                                                                                                                                                                                                                                                                                 GGQGAGAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQG 370
                                                                                                                                                                                                                                                                                                                                                                                     497
                                                                                                                                                                                                                                                                                                                                                                                                                                                 587
                                                                                                                                                                         Gaps
                                                                                                                                                                                                26
                                                                                                                                                                                   SQG-AGRGGLGGQGAGAAAAAAAAGGAGQGGLGGQGAGQGAGAGAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                  GGYGGLGSQGAGR---GGQGAGAAAAAVGAGQEGIRGQGAGQGGYGGLGSQG-SGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                              GGGGAGAAAAAAGGAGGGGGGGGGGGGGAGAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                                                                                                                           QGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                   AGAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                    LGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT-----SGP
                                                                                                                                                                                                                                                                                                                                                                                                            GGYGP-GQQTSGRGGLGGQGAGAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                    GGQGAGAAAA-------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                       39; Indels 176;
                                                                                                                                                            Length 747;
                                                                                                                                               850E44B0D649E012 CRC64;
                                                                                                                                                            61.1%; Score 2135; DB 1;
68.3%; Pred. No. 7.9e-89;
                                                                                                                                                                        15; Mismatches
                                                                                                                                          (IN REF.
                                         114.
115.
117.
119.
                                                                                                                                               60528 MW;
                                                                                                                                                                       495; Conservative
1191
2204
2305
2305
2305
2305
2305
244
247
247
247
247
247
                                                                                                                                               747 AA;
                                                                                                                                                                 Similarity
REPEAT
REPEAT
CONFLICT
                                                                                                                        CONFLICT
                                                                                                                                               SEQUENCE
                                                                                                                                                           Query Match
                                                                                                                                                                  Local
                                         REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                   REPEAT
REPEAT
REPEAT
                 REPEAT
REPEAT
REPEAT
REPEAT
            REPEAT
                                                                                                                                                                 Best Loca
Matches
                                                                                                                                                                                                                                                           200
                                                                                                                                                                                                                                    140
                                                                                                                                                                                                                                                                                               202
                                                                                                                                                                                                                                                                                                           290
                                                                                                                                                                                                                                                                                                                       261
                                                                                                                                                                                                                                                                                                                                   349
                                                                                                                                                                                                                                                                                                                                               312
                                                                                                                                                                                                                                                                                                                                                           391
                                                                                                                                                                                                                                                                                                                                                                                   451
                                                                                                                                                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                                                                                                                    551
                                                                                                                                                                                                                                                                                   251
                                                                                                                                                                                                                                                                                                                                                                       371
                                                                                                                                                                                                                                                                                                                                                                                                            498
                                                                                                                                                                                                                                                                                                                                                                                                                       476
                                                                                                                                                                                                                                                                                                                                                                                                                                               532
ద
                                                                                                                                                                                                                       g
                                                                                                                                                                                    ð
                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               39;
                                                          989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             s:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----TSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGR------GGLGG 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mycobacterium tuberculosis.
Bacteria, Actinobacteria; Actinobacteria (class); Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-H37RV;

X MEDLINE-98395987; PubMed-9634230;

A Goldon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,

A Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,

Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

Davies R., Deviin K., Feltwell T., Gentles S., Hamilin N., Holroyd S.,

A Hornsby T., Jaglas K., Krogh A., McLean J., Moule S., Murphy L.,

A Hornsby T., Jaglas K., Krogh A., McLean J., Moule S., Murphy L.,

A Rutter S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

"Deciphering the blology of Mycobacterium tuberculosis from the complete genome sequence.";

Nature 393:537-44(1998)

C -- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Harris
<u> QGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGGAAAAAAAAAAAGGAGQGGY</u>
                             Holroyd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 GRGGLGG----QGAGAAA------AAAAAAGG-AGQGGYGGLGSQG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               235;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            36.6%; Score 1277.5; DB 1; Length 1901; 39.7%; Pred. No. 1.1e-50; Live 28; Mismatches 270; Indels 235;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYPOTHETICAL PE-PGRS FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C7B1923D5D0146CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                   30-MAY-2000 (Rel. 39, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv3508 precursor.
RV3508 OR MTV023.15.
```

m

```
1006 DGASGLGLGLSGFDGGQGGGGGGGAGGSNGAGGAGGNGGDGGDGATGAAGLGDNGG 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                             215
                                                                                                                         GOGGYGGLGSQ----GTSGPGGYGPGQQTSGRGGLGGQGA------GAAAAAA 360
                                                                                                                                                                                                        391
                                                                                                                                                                                                                                                                                                                                                                                                         587
                                                                                                                                                                                                                                                                                                                                                                                                                                                588 LGCQ-GAGAAA----TSGRGCLGGQ 629
       412 GTGFAGGAGGVGGGGRAIAGGINGSGGAGGTGGOGGAGGMGGSGADNASGIGADGGAGG 471
                                                531
                                                                                        532 FSGGAGGAGGATGVGGTNGSGGQG-GTGGAGGAGGAGGVGADNPTGIGGTGGTGGKG 590
                                                                                                             216 TSGRGGLGGQG--AGAAAAAAAAAGG-AGQGGYGGLGSQGT---SGRGGLGGQGAGAAA 269
                                                                                                                                                     270 AAAAAAAAGGGGYGG-----LGSQGTSGRGGLGGQ-GAGAAAAAAAAAAAGGA---- 317
                                                                                                                                                                                                                                                                                                  829 GSAGAGGINGAGGAGGNGGDGGDGATGAAGLGDNGGVGGDGGAGGAAGNGGNAGVGLTAK 888
                                                                                                                                                                                                                                                                                                                      889 AGDGGAAGNGGAGGAGGAGGAGDNNFNGGQGGAGGQGGGGGGGGGAGTTSINANGGAGGNG 948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hexapoda;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Merazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda
Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Bombycoidea; Bombycidae; Bombyx.
                                         ς;
:
                             QG--AGAAAAAAAAAGG-------QGGGYGGYGGLGSQGT-----
                                                                                                                                                                                                                                                                                                                                                               482 GQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGGYGGLGS
                                                                     ----SGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAGGAGQ-----GGYGGLGSQG
                                                                                                                                                                                                                                      361 AAAAGGA-GQGGYGGL-----GSQGTSGRGGLGGQGA-------
                                                                                                                                                                                                                                                           769 AGGAGGATGTGGTGGVVGATGSAGIGGAGGRGGDGGDGASGLGLGLSGFDGQGGQGGAG
                                                                                                                                                                                                                                                                                                                                                                                                        542 QGTS-----GRGGLGGQGAGAAAAAAAAAAGGAGGGGGGGGGGGGTSG---RGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Esnault C.,
Li Z.-G.;
gene.";
                                                                                                                                                                                                                                                                                GAAAAAAAAAGGA-GQGGYGGLGSQGTSG---RGGLGGQ-GAGAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-20330362; PubMed-10871375; Zhou C. Z., Confalonieri F., Medina N., Zivanovic Y., Yang T., Jacquet M., Panin J., Duquet M., Perasso R., Fithe organization of Bombyx mort fibroin heavy chain Nucleic Acids Res. 28:2413-2419(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GA----GAAAAAAAAAAGGA----GQGGYGGLGSQGTSGPGGYG 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FBOH_BOMMO STANDARD; PRT; 5263 AA. P05790; 026379; 017220; 017220; 01700-1988 (Rel. 09, Created) 16-0CT-2001 (Rel. 40, Last sequence update) 16-0CT-2001 (Rel. 40, Last annotation update) Fibroin heavy chain precursor (Fib-H) (H-fibroin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bombyx mori (Silk moth)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-7091;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FERCH_BOMMO
ID FEBOH_BK
DT 01-NOV
DT 01-NOV
DT 16-OCT
DE FIBOT
GN FIBH.
GN BOMBYX
OC EURAFY
OC INSECT
OC DILTYS
OX NCBI_T
RN (1]
RN (1]
RX MEDLI)
RA Zhou (
RT FIBE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1126
                                                                     167
                                                                                                                                 591
                                                                                                                                                                         649
                                                                                                                                                                                             318
                                                                                                                                                                                                                  709
                                                                                                                                                                                                                                                                               392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 3
       g
                                              g
                                                                                        ద
                                                                                                                             g
                                                                                                                                                    δ
                                                                                                                                                                   q
                                                                                                                                                                                        δ
                                                                                                                                                                                                               q
                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                            ò
                                                                    ò
                                                                                                          οŽ
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SECTION OF SILK GLANDS.
-1- DOMAIN: COMPOSED OF ANTIPARALLEL BETA SHEETS. THE STRANDS OF THE BETA SHEETS. THE STRANDS OF SILK FIBROIN ARE COMPOSED OF MICROCRYSTALLINE ARRAYS OF (-GILY-SER-GLY-ALA-GLY-ALA-)N INTERRUPTED BY REGIONS CONTAINING BULKIER RESIDUES. THE FIBER IS COMPOSED OF MICROCRYSTALLINE ARRAYS ALTERNATING WITH AMORPHOUS REGIONS.
[2]
SEQUENCE OF 1-168 FROM N.A.
MEDLINE-80045039; PubMed-498286;
MEDLINE-80045039; PubMed-498286;
Taulimoto Y., Suzuki Y.;
"The DNA sequence of Bombyx mori fibroin gene including the 5'
flanking, mRNA coding, entire intervening and fibroin protein coding regions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PARTIAL SEQUENCE FROM N.A. MEDIAL SEQUENCE FROM N.A. MEDILINE-9435842; Pubmed-7916056; MAITA K., ICHIMURA S., James T.C.; "Highly repetitive structure and its organization of the silk fibroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Kinshu X Showa;
MEDLINE-89094868; PubMed=3210244;
Mita K., Ichimura S., Zama M., James T.C.;
Mita K., Ichimura S., Zama M., James T.C.;
Specific codon usage pattern and its implications on the secondary structure of silk fibroin mRNA.";
J. Mol. Biol. 203:917-925(1988).
                                                                                                                                                                                                                                              PARTIAL SEQUENCE FROM N.A.
MEDLINE-19211211; PubMed-455439;
TSUJIMOLO Y., SUZUKI Y.;
"Structural analysis of the fibroin gene at the 5' end surrounding redions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 5179-5263 FROM N.A., AND DISULFIDE BONDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FIBROIN HEAVY CHAIN.
HIGHLY REPETITIVE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AF226688; AAF76983.1; EMBL; V00094; CAA23432.1; EMBL; V00097; CAA23433.1; EMBL; S74439; AAB31861.1; EMBL; X13869; CAA32076.1; EMBL; M35378; AAA7839.1; PIR; S01844; S01884.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-99296390; PubMed-10366732;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mol. Evol. 38:583-592(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                       PARTIAL SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21
5263
5206
                                                                                                                                                                                                                                                                                                                                                                                    Cell 16:425-436(1979).
                                                                                                                                                                                                                                                                                                                                                              surrounding regions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1
22
149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-J-139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIGNAL
CHAIN
DOMAIN
```

4

```
26;
                                                                                                                                                                4532 FGPYVANGGYSGYEYAWSSESDFGTGSGAGAGSGAGAGSGAGAGYGAGYGAGYGAG
                                                                                                                                                                                                                                                                                                     4652 AGYGAGYGAGAGAGYGAGAGVGYGAGAGYGAGAGSGAGSGAGAGSGSGAGAGSGSGAGAG
                                                                                                                                                                                                                                                                                                                                                                                                       5010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4352 ASGAGAGSGAGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGA 4411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QGTSGRGGLGGQGAGAAAAAAAAAAAAAGGGG-GYG-GLGSQGTSGPG---GYGPGQQTSG 180
                                                                                                                                                                                                                                             -----RGGLG------GQGAGAAAAAAAAAAAGAGGG------GYG-- 209
                                                                                                                                                                                                                                                                                        ---GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAG 266
                                                                                                                                                                                                                                                                                                                                       294
                                                                                                                                                                                                                                                                                                                                                                                295 RGGLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGPG-GYGPGQQTSGRGGLG---- 349
                                                                                                                                                                                                                                                                                                                                                                                                                                404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                664
                                                                                                                                                 74 AGGAGQG-----GYG---GLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGS 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AGQGGYGGLGSQGTSGRG-----GLG-GQGAGAAAAAAAAAGGGGGGGGGGGGGGGT 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SGRGGLGGQGAGAAAAAAAAAGGAGQG------GYG-GLGSQCTSGPG---GYG 501
                                                                                 Gaps
                                                                                                      AAAAAAAAAAAGGAGGG-----GYG-----GYG-------GLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                              PGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAAAAGGAGQG-GYG-GLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GYG-GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQG----GYG-GLGSQGTSGPG-GY
                                                                               49; Mismatches 296; Indels 130;
                                                        Length 5263;
 INTERCHAIN (WITH LIGHT CHAIN)
                     10 10 C -> V (IN REF. 2).
5263 AA; 391586 MW; 8EE11D3AOA47440E CRC64;
                                                        33.4%; Score 1168; DB 1; 39.7%; Pred. No. 1.4e-45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           627 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                               Matches 313; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  32,
32,
41,
                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5131 GAGTGSSG 5138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (Rel.
(Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GPGQQTSG 672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1995
01-NOV-1995
15-JUN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SPD2_NEPCL
P46804;
                                 SEQUENCE
          DISULFID
                                                        Query Match
DISULFID
                        CONFLICT
                                                                     Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 4
SPD2_NEPCL
                                                                                                                                                                                               126
                                                                                                                                                                                                                     4472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5071
                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                        210
                                                                                                                                                                                                                                                                                                                                     267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      502
FF 50
                                                                                                      ð
                                                                                                                                                  ð
                                                                                                                                                                       g
                                                                                                                                                                                               õ
                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                윱
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           44444
4444
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outsitution the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36;
                                                                                                                                                                                       240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       77 AGQGGYGGLG-SQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQ----GTSGR 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SGPGSAAAAAAASGPGQQGPGGYGP-GQQGPGGYGPGQQGPSGPGSAAAAAAASGPGQ 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGQGGYGGLGSQGTSGRGGLG-GQ----GAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGR 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GCLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -GQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG----GLGGQGAGAAAAAAAAAAAGG
                                                       Araneae;
Nephila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPROXIMATE TANDEM REPEATS
Spidroin 2 (Dragline silk fibroin 2) (Fragment).
Nephila clavipes (Orb spider).
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CB9B63779B2C594B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           32; Mismatches 180;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           32.5%; Score 1135; DB 1; 47.5%; Pred. No. 1.2e-44;
                                                                                                                                                                          MEDLINE=92406876; PubMed=1527052;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54184 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M92913; AAA29381.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 324; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       516 :
627 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                    NCBI_TaxID=6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Silk; Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
REPEAT
REPEAT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241
     SOURCE STATE STATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
```

= -= -

= = - Ŋ

Page

```
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                         410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                s:
                                                        468
                                                                                                                                                                                            648
                                                                                                                                                                                                                                                                                                                                                                                                         GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGA 355
                                                                                                                                      311 AAAAAAAAGPGQQGLGGYGP-GQQGPGGYGPGQQGPGGYGPGSASAAAAAAGPGQQGPGGY 369
                                                                                                                                                                                                                                                                                                                  GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGCL 588
                                                                                                                                                                                                                                                                                                                                                  --AGPGGYGP-AQQGPSGPG------IAASAASAGPGGYG------PAQQGP 500
QGPGGYG----PGQQGPGGYGPGQQGLSGPGSAAAAAAAGPGQQGPGGYGP-GQQGPSGP 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ö.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Harris
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-9829597; PubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris Cole S.T., Brosch R., Parkhill J., Gas S., Bary C.E. III. Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd Burnsby T., Jagels K., Kroph A., McLean J., Moule S., Murphy L., Oliver S., Seeger K., Skelton S., Squares S., Squares T., Salston J.S., Taylor K., Mhitchead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
                                                                                                                                                                                                                                             356 AAAAAAAAGGAGQGGYGGLGSQGTSGRG----GLGGQGAGAAAAAAAAA-GGAGQGGY
                                                                                                                                                                        GGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSG--RGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                     GGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Whole genome comparison of Mycobacterium tuberculosis clinical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -1 - CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              914 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WAG22 antigen precursor.
WAG22 OR RV1759C OR WT1807 OR MTCY28.25C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        GYGGLGSQ-----GISGPG 662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        556 SSAALSSVISNAVSQIGASNPG 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-MAY-2000 (Rel. 39, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-CDC 1551 / Oshkosh;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WA22_MYCTU
006794;
                                                                                                                                                                                                          370
                                                                                                                                                                                                                                             469
                                                                                                                                                                                                                                                                                                                529
                                                                                                                                                                                                                                                                                                                                                  464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
                                                             셤
                                                                                                                                    셤
                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                      ò
                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                            ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                          32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          176 TANNGIGGAGGAGGSAMLFGAGGAGGAGGAATSLVGGIGGTGGTGGTAGMLAGAAGAGGA 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RGGLGGQG--AGAAAAAAAAAAGGAGQGGYGGLGSQG--TSGRGGLGGQGAGA-----A 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAGGAGOGGY - GGLGSOGTSGRGGLGG - - - OGAGAAAAAAAAAAGGGGGGGGGGGGO 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     643 GLGGCGGGAFTAGVTTGGAGGTGGAAGLFANGGAGGAGGTGSTAGGAGGAGGAGGLYAHG 702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              821
                                                                                                                                                                                                                                                                                                                                                                                                                                                        GNGANGAPGTGANGGDAGWLIGNGGAGGSGAKGANGGAGGPGGAAGLFGNGGAGGAGGTA 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66 AAAAAAAAAAGGA-----GQGGYGGLGSQGTS---GRGGLGGQGAGAAAAAAAAAAAGGA 115
                                                                                                                                                                                                                                                                                                                                                                                                                        GSMASGRGGLGGQGAGAAAAAAAAAGGAG----QGGYGGL-GSQGTSGRGGLGGQGAGA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGFSFSTAGGAGGAGGAGGLFTTGGVGGAGGQGHTGGAGGAGGAGGLFGAGGMGGAGGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             160 GLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAGG-----AGQGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLGSQGT----SGRGGLGGQGAGAAAAAAAAAGGAG-----QGGYGGLGS-----QGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   174 EGGIGALAGKSGFGGFGGLLLGADGYNAPESTSPWHNLQQDILSFINEPTEALTGRPLIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -GSQGTSGPG-GYGPGQQTSGRGCLGGQGAGAAAAAAAAAAGGA-----GQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --GRGGLGGQ-----GAGAAAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAG--QGGYGGLGSQGTS----
                                                                                                                                                                                                                                                                                                                                                                                         304; Indels 159;
                                                                                                                                                                                                                                                                                                                                                         DB 1; Length 914;
                                                                                                                                                                                                                                                                                        1 30 POTENTIAL.
31 914 WAG22 ANTIGEN.
914 AA; 74354 MW; F6953C3DBEBE6AC8 CRC64;
                                                                                                                                                                                                                                                                                                                                                       31.7%; Score 1106.5; DB 38.2%; Pred. No. 2.7e-43; ative 31; Mismatches 304
                                                                                                                                                                                                                                                                   Antigen; Repeat; Signal; Complete proteome. SIGNAL 1 30 POTENTIAL.
                                                                                                                                                                 EMBL; Z95890; CAB09322.1; -. EMBL; AE007040; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAAAAAGGAGGGYGGL------
                                                                                                                                                                                                  TIGR; MT1807; 'Tubercutist; Rv1759c; 'Tubercutist; Rv1759c; 'LuterPro: IPROGO084; PE_region. Pfam; PF00934; PE; 1.
 FRAMESHIFT IN POSITION 85.
                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                         Matches 306;
                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                        116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     379
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

us-09-490-291-4.rsp

41.38;

Local Similarity

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation-the European Bloinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lab.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Fleischmann R.D., Alland D., Elsen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weldman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C STRAIN-H37RV;

WEDLINE-9895987; PubMed-9634230;

WEDLINE-9895987; PubMed-9634230;

A Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., A Gordon S.Y., Eiglumeier K., Gas S., Barry C.E. III, Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Badcock K., Beltwell T., Gentles S., Hamlin N., Holroyd S., A Davles R., Jagels K., Krogh A., McLenn J., Woule S., Murphy L., Oliver S., Osborne J., Quall M.A., Rajandream M.A., Rogers J., A Butten J.E., Taylor K., Whitehead S., Barrell B.G.;

The State of the biology of Mycobacterium tuberculosis from the R. Nature 393:537-544(1998).
----AGAAAAAAAAAAAGGAGGGYGG-LGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQ 647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         s.
                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome comparison of Mycobacterium tuberculosis clinical and
                                    822 RSGFGGAGGAGGASDTGAGGNGGAGGKAGLYGNGGDGGAGGDGATSGKGGAGGNAVVIGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; ALL. 1

I TIGR: MT2712; -...
R TUBECLULST; RV2634c; -...
R TUBECLULST; RV2634c; -...
JR InterPro; PR000084; PE_region.
DR Probom; PD001223; PE_region; 1.
DR Probom; PD001223; PE_region; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 51 51 V -> L (IN REF. 2).
FT CONFLICT 63 63 Q -> H (IN REF. 2).
TALL 174 A -> T (IN REF. 2).
                                                                                                                                                                                                                                                                            01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv2634c.
MyCobacterium tuberculosis.
                                                                                                                                                                                                                                        778 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; Z80225; CAB02341.1; -.
EMBL; AE007103; AAK47026.1; ALT_INIT.
                                                                                                               882 GGNGGNAGKAGGTAGAGGAG 901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
STRAIN-CDC 1551 / Oshkosh;
                                                                                      GGYGGLGSQ - - GTSGPGGYG
                                                                                                                                                                                                                                          STANDARD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBFAMILY.
                                                                                                                                                                                                                                        YQ34_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Whole
ð
                                         q
                                                                                                                           ద
```

DB 1; Length 778;

```
34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       406
                                                                                                                                                                                                                                                                                                                                                                  410
                                                                                                                                                                                                                                                                                                                                287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         135 GGLLLGNGGNGGSGAPGQPGCAGGDAGLIGNGGTGGKGGDGLVGSGAAGGVGGRGGWLLG 194
                                                                                                                     GSQGTSGR-GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRG 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SGRGGLGGQ--GAGAAA----AAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       s:
                  99; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
STRAIN-H37RV;
MDLINE-89825987; PubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris Gordon S.V., Elglmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd
                                                                                                                                                                                       180 GRGCLGGGGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAG
                                                  11 GSMASGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQ------GTSGRGG--LG
                                                                                                                                                       195 NGCIGGAGGAAGAILVG-GIGGVG--GAIGLIGSGGFG--GAGGAAAGVGTIGGVGGSGG
                                                                                                                                                                                                                                                                                   GAGQG-----GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAG------GAGQGGYGGL
                                                                                                                                                                                                                                                                                                                                                               351 AGGNGGNANWFGSGCAGCGGGTGLAGTNGVNPGSIANPNTGANGTDNSGNGNQTGGNGGP
                                                                                                                                                                                                                                                                                                                                                                                                                                   411 GPAGGVGEAGGVGGOG-----GLGESLDGNDGTGGKGGAG--GTAGTDG-GAG----GAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          518 GDGNGALGAAG--GNGGTGGAGGNGGRGGMLIGNGGAGGAGGTGGTGGGGAAGFAGGVGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           576 AGGEGLIDGAGTAEGGT -- GGLGGLG -- GVGGTGGMG-----GSGGVGGNGGAAGSLIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 407 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGA-GQGGYGGLGSQGTSG-RGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             525 AAAAGGA-GQGGYGGLGSQGTSGRGGLGGQ-GAGAAAAAAAAAAAGGAGQGGYGGLGSQGT
                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv0278c precursor.
RV0278C OR MT0291 OR MTV035.06C.
; Pred. No. 1.4e-40; 21; Mismatches 290;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 636 AAAAAAAGGAGGGYGG----LGSQGTSGPGGYGPGQQ 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   740 GGDFALGGNGGAGGAFGSSGIQGNMGP----PGTQ 774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           957 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mycobacterium tuberculosis
                Matches 289; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Y278_MYCTU
P56877;
                                                                                                                                                                                                                                                                                                                              240
                                                                                                                                                                                                                                                                                                                                                                                                 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                583
                                                                                                                                                   g
                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                     à
                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
```

~

```
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-1773;
                                                                                                                                                                                                                                                                                                                                                                                                    675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                 GCRR 882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-H37RV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Y747_MYCTU
O53810;
                                                                                                                                                                                                                                                                                                                                                                                                      GIRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Y747_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                 879
                                                                                                                                                                                                                                                    719
                                                                                                                                                                                                                                                                                  563
                                                                                                       412
                                                                                                                                    009
                                                                                                                                                                 991
                                                                                                                                                                                             629
                                                                                                                                                                                                                          514
                                                                                                                                                                                                                                                                                                             777
                                                                                                                                                                                                                                                                                                                                            613
                                                                                                                                                                                                                                                                                                                                                                       835
                                                                                                                                                                                                                                                                                                                                                                                                    67.3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                   셤
                                                                                                                                                                                        q
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                       ð
                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                          å
                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics. The There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31;
                                                                                                                              STRAIN-CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Oimayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        370 GAGGAGGNAGLLFGSGGSGGAGGFGF-ADGGQGGPGGNAGTVFGSGGAGGNGGVGQGFAG 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         132 GGWLIGNGGAGGSGAAGVNGGAGGNGGAGGNGGAGGLIGNGGAGGAGGVASSGIGGSGGA 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAAAA--GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAG-----QGGY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  192 GGNAMLFGAGGAGGAGGGVVALTGGAGGAGGAGGNAGLLFGAAGVGGAGGFTNGSALGGA 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            252 GGAGGAGGLFATGGVGGSGGAGSSGGAGGAGGLFGAGGTGGHGGFADSSFGGVGGAG 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GAGGLFGAGGEGGSGGHSLVAGGDGGAGGNAGMLALGAAGGAGGIGGDGGTLTAGGIG-- 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           286 GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGGGGGGGGGG----LGSQGTSGPGGYGPGQQ 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---GAGAAAAAA 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                       Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GSMASGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGL-GSQGTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121 GGLGSQG----TSGRGGLGGQGAGAAAAAAAAGGAGGGGYGGLGSQGTSGPGGYGPGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       177 QTSGRGGLGGQGAGAAAAAAAAAGGAG-----QGGYGGLGSQ-GTSGRGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 226 GAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GIG --GAGGTPGLIGNGGNGGNGGASAVTG --GNGGIGGTGVLIGNGGNGGSGGIGAGK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HYPOTHETICAL PE-PGRS FAMILY PROTEIN
   Ľ.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; "Dechpering the biology of Mycobacterium tuberculosis from the complete genome sequence."; Nature 393:337-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 29.4%; Score 1028; DB 1; Length 957; Similarity 37.4%; Pred. No. 8.3e-40; Conservative 29; Mismatches 310; Indels 152;
                                                                                                                                                                                                                                    laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000084; PE_region.
Pfam: PF00934; PE: 1.
Hypothetical protein; Repeat; Signal; Complete proteome.
SIGNAL 1 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              M -> I (IN REF. 2).
MISSING (IN REF. 2).
R -> G (IN REF. 2).
71EBABD417FBA47C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AL021930; CAA17353.1; -.
EMBL; AE006936; AAK44511.1; ALT_INIT.
TIGR; MT0291; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RV0278C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40 M
163 M
807 R
81905 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     342 TSGRGGLGGQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 293; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tuberculist; Rv0278c;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40
158
807
957 AA;
                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                 SUBFAMILY.
                                                                                                                                                                                                         Bishai W.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query
```

ð 셤 à g g

δ

à

g ŏ 8 ò g

```
STRAIN=CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ς.
-AGVGGVSGLLLGLDGFNAPASTSPLHTLQQNVLNVVNEPFQTLTGRPLIGNGANGTPGT 542
                                                                                                                                                                            543 GADGGAG-GWLFGNGANGTPGTGAAGGAGGWLFGNGGNGGHGATNTAATATGGA--GGAG 599
                                                                                                                                                                                                                                                                     GLGSQGTSGRGGLGGQGAGAAAAAAAAGGA----GQGGYGGLG--SQGTSGRGGLGGQ 465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GGDAGLLFGAAGTGGHGAAGGVPAGVGGAGGNGGLFANGGAGGAGGFNAAGGNGGLF 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GOGAGAAAAAAAAAAGGAGOGGYGGLGSQGTSGRGGLGGQG------AGAAAAAA 562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GTGGTGGAGINFGAGGNGGNGGLFGAG--GTGGAAGSGGSGITTGGGGHGGNAGLLSLGA 776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAAAGGAG-----QGGYGGLGSQGT----SGRGGLGGQGAGAAAAAAAAAGGAGQGGY 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQ-GTSGPGGYGPGQQTS 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Harris
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-9829587; PubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris Gordon S.T., Brosch R., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamiln N., Holroyd Horsby T., Jagels K., Kroph A., McLean J., Moule S., Murphy L., Oliver S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor R., Whitehead S., Barrell B.G.;
"Deciphering the blology of Mycobacterium tuberculosis from the complete genome sequence."
                                                                                                                                                                                                                                                                                                                                                              GI-LFGTGGNGGTGGIATGAGGIGGAGGAGGVSLLIGSGGTGGNGGNSIGVAGIGGAGGR
                                                                                        AAAGGAGGGYGGLGSQGTSGRG-----GLGGQGAGAAAAAAAAAAAAGGAGGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                       GA-----GAAAAAAAAAAGG--AGQGGYGGLG----SQGTSGPGGYGPGQQTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SGGAGGSGGASSLAGGAGGTGGNGALLFGFRGAGGAGGHGGAALTSIQQGGAGGA--GGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Whole genome comparison of Mycobacterium tuberculosis clinical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv0747 precursor.
Rv0747 OR MT0772.5 OR MTV041.21.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   801 AA.
```

us-09-490-291-4.rsp

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                   31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83 GGGAYAAAEAAAVSPLLAPINAQFVAATGRPLIGNGANGAPGTGANGGPGGWLIGNGGAG 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLGSQG-TSGRGGLGGQGAGAAAAAAAAAAGGAGQ--GGYGGLG-----SQGTSGPGGY 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                68 aaaaaaagaagggyggl-gsqgtsg-----rgglggggaaaaaaaaaaagagggyg 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               320 NGGASTGGGDGGPGGAGGTGVLIGNGGNGGSGGTGA--TLGKAGIGGTGGVLLGLDGFTA 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   657
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---AAAGGAGGGYGGLGSQGTSGRGGLGGQ--GAGAAA 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     201 GFSNGGATGGAGGAG-GLFGAGRERGSGGSGNLTGGAGGAGGNAGTLATGDGGAGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     378 PASTSPLHTLOODVINNVNDPFOTLTGRPLIGNGANGTPGTGADGGAGGWLFGNGGNGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GTSGRGGLGCQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               G----AGQGGYGGLGSQ-GTSGRGCLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GPGQQTSGRGGLGGQGAGAAAAAAAAAGGAG------QGGYGG----LGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RGGL----CGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----GAAAAAAAAAAAAAAGGGG--GYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GTI--GGVNG-GAGGAGGALFGTGGTGGSGGPGATGLGGIGGAGGAALLFGSGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SGGAGAVGGNGGAGGNAGALLGAAGAGGAGGAGAVG--GNGGAGGNG-GLFANGGAGGPG
                                                                                                                                                                                                                                       HYPOTHETICAL PE-PGRS FAMILY PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                  33; Mismatches 289; Indels 144;
                                                                                                                                                                                                                                                                                                                                                                                                                      Score 988; DB 1; Length 801;
Pred. No. 4.4e-38;
                                                                                                                                                                                                                                                              G -> 5 (IN REF. 2).
R -> G (IN REF. 2).
R -> G (IN REF. 2).
K -> R (IN REF. 2).
S -> G (IN REF. 2).
T -> I (IN REF. 2).
T -> I (IN REF. 2).
T -> P (IN REF. 2).
T -> P (IN REF. 2).
MISSING (IN REF. 2).
W; EA54C9Bf45A00F41 CRC64;
                                                                                                                                                                                                            Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                     4.4e-38;
                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                  TIGR; MT0772.5; -.
TubercuList; Rv0747; -.
InterPro; IRR00084; PE_region.
Pfam; PF00934; PE; 1.
Hypothetical protein; Repeat; S
SIGNAL 1
                                                                                                                                       AE006968; AAK45011.1; -.
                                                                                                                          EMBL; AL021958; CAA17514.1; -.
                                                                                                                                                                                                                                                                                                                                                                                           65407 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                      28.38;
37.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 GQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 284; Conservative
                                                                                                                                                                                                                         30
                                                                                                                                                                                                                                                                                                         295
300
338
                                                                                                                                                                                                                                                                                                                                                                                         801 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                 188
225
227
227
230
300
338
                                                                                                                                                                                                                                                                                                                                                               CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                              CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                     CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                      Duery Match
                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                      Local
                                                                                                                                                                                                                                       CHAIN
                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               104
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and flor statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genomics 23:125-131(1994).

-I-FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND

-I-FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND

NUCHAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.

-I-SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER

INTO AN EXTRENSIBLE 3D NETWORK.

-I-SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.

-I-PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.
                                                                                                  714 DGGAG--GKSGAFGLGGDGGAGGATGLSGAFHIGGKGGVG--GSAVLIGNGGNGGNG 769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LGSQGTSGRGGL-GGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    106 AAAAAAG-GAG--QGGYGGLGSQGTSGRGLGGQGA------GAAAAAAAA 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ဌ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-BALB/C; TISSUE-Lung;
MEDLINE-95130069; PubMed-7829060;
Wydner K.S., Sechler J.L., Boyd C.D., Passmore H.C.;
"Use of an intron polymorphism to localize the tropoclastin gene t mouse chromosome 5 in a region of linkage conservation with human
 ---GAGGLLFGSGGAGGSGGFSNSGNGGAGGAGGAGGLLVGSGGAGGAGASAT-GAATGG
                                                                       574 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----AGAAAAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        36; Mismatches 249; Indels 214;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 860;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  71955 MW; OCOBESAAE1EDD7F1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRINTS; PRO1500; TROPOELASTIN.
Structural protein; Repeat; Signal; Connective tissue.
STGNAL. 1 27 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27.9%; Score 976; DB 1; 39.3%; Pred. No. 1.6e-37;
                                                                                                                                                                                                                                                                                                              01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
01-0CT-1996 (Rel. 34, Last annotation update)
                                                                                                                                                                                                                                                                               860 AA
                                                                                                                                                                  770 GNAGKSGGAPGPSGAGGAGGLLGENGLNG 799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ELASTIN
                                                                                                                                               634 A-AAAAAAAGGAGQGGYGG--LGSQGTSG
                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR003979; tropoelastin.
                                                                                                                                                                                                                                                                                                                                                          01-OCT-1996 (Rel. 34, Last annot:
Elastin precursor (Tropoelastin)
                                                                                                                                                                                                                                                                                                                                                                                                                              Chordata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; U08210; AAA80155.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 323; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11 GSMASGRGGLGGQG----
                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     860
                                                                                                                                                                                                                                                                                                                                                                                                               Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MGI:95317; Eln.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    860 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                             ELS_MOUSE
P54320;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local
519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          47
                                  658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             138
                                g
                                                                                                                                                                                                                                                                                              à
                                                                                                      g
                                                                                                                                                 à
                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

σ

```
235
                                                                                                                                                                                                                                                                                                                                                                                     476
                                                                                                                                                                                                                                                                                                                                                                                                                   648
                                                                                                                                                                                                                                                                                                                                                                                                                                               516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      572
                                                                                                                         305 GTGVGSQAAAAAAKAAKYGAGGAGVLPGVGGGGIPGGAGAIPGIGGIAGAGTPAAAAAK 364
                                                                                                                                                          -----GQGGYGGLGSQGTSGRGGLGGQGAG 266
                                                                                                                                                                            AAAAAAAAAAAGGA----GLGSQGTG----GLGSQGTSGRGLGGQGAGAAAAAAAAGGAGQ 319
                                                                                                                                                                                                                                                                        GGYGGLGSQGTSGPG---GYGPG--QQTSGRGLGGQGAGAAAAAAAAAA----AG-GAG 368
                                                                                                                                                                                                                                                                                                    544
                                                                                                                                                                                                                                                                                                                                                         545 VGGVPGGVGVGGIPGGVGVGG------VPGGVGPGGVTGIG----AGPGGLGGA 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                              753
304
                                                                                                                                                                                                                               QGGY-GGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQ 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chórdata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cDNA cloning.";
                                                                                                                                                                                                                                                                                                                                                                                                      AGGAGQG--GYGGLGSQ-----GTSGPGGYG----PGQQTSGRGGLGGQ----
                                                                                                --GAGAAAAAAAAAGGAGQGGYG---GLGSQGTSGR-----GGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                 485 YGAGGAGALGGLVPGAVPGALPGAVPAVPCAGGVPGAGTPAAAAAAAAAAAAKAGLGPG
                                                                                                                                                                                                                                                                                                                                                                                     GAGAAAAAAA-----AAAAG-GAGQGGYG-GLGSQGTSGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                            477 AAG----GAG-----QGGYGGLGSQ-GTSGPGGYGPGQQTSGRGGLGGQG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      573 GY--GGLGSQGTSGRGGLG-GQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQ
                                                                     GGGAFSGIPGVGPFGGQQPGVPLGYPIKAPKLPGGYGLPYTNGKLPYGVAGAGGKAGYPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 781-864 FROM N.A.
MEDLINE-88330868; PubMed=2971041;
Deak S.B., Pierce R.A., Belsky S.A., Riley D.J., Boyd C.D.;
"Rat tropoelastin is synthesized from a 3.5-kilobase mRNA.";
J. Biol. Chem. 263:13504-13507(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING. MEDLINE-91104868; Pubmed-1702999; Plerce R.A., Deak S.B., Stolle C.A., Boyd C.D.; "Heterogeneity of rat tropoelastin mRNA revealed by Biochemistry 29:9677-9683(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      811 GFGLSPIYPGGGAGGLGVGGKPPKPYGGALGALGYQGGCFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  630 GAGAAAAAAAAAGGAGQGG----YGG-LGSQGTSGPGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-OCT-1996 (Rel. 34, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Elastin precursor (Tropoelastin) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-OCT-1996 (Rel. 34, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rattus norvegicus (Rat).
                                                                                                                                                        AAAGGA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ELS_RAT
099372;
                                                                     245
                                                                                                                                                        236
              185
                                         150
                                                                                                                                                                                                                                           425
                                                                                                                                                                                                                                                                        320
                                                                                                188
                                                                                                                                                                                                               267
                                                                                                                                                                                                                                                                                                                                                                                     428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                    셤
                                                                                                ò
                                                                                                                         g
                                                                                                                                                 δλ
                                                                                                                                                                               셤
                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                             å
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                    ογ
                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            엄
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              38;
                                                                                                                                                                                              -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.
                            MEDLINE-92241859; PubMed-1572637;
Pierce R.A., Alatawi A., Deak S.B., Boyd C.D.;
Generics of the rat tropoelastin gene associated with alternative splicing.";
Generics STRUCTURN: PROTEIN OF TISSUES SUCH AS AORTA AND INCHAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOMPLETELY.
-1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAAAAAAAGGGGGGGGGGGGGT-----GSQG-----TSGRGGLGGQGAGAAAAAA 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             173 IGAGVGAGGKPGKVPGVGLPGVYPGGVLPGTGARFPGVGVLPGVPTGTGVKAKVPGGGGG 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                53 GGLGGGALGPGGKPPKPGAGLLGAFGAGPGGLGGAGPGAGLSYASRPGGVLVPGGGAGAA 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAYKAAAKAGAGLGGIGGVPGGVGVGGVPGAVGVGGVPGAVGGIGGIGGLGVSTGAVVPQ 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99
                  OF 264-533 AND 558-864 FROM N.A., AND ALTERNATIVE SPLICING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AGGAGQGGYGGLGSQGTSGR-GGLGGQGAGAAAAAAAAAAGG-AGQGGYGGLGSQGTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 GGLGGGGAGAAAAAAAAAAAG-----GAGGGGYGGLGSQG----TSGRGG--LGGQGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           110 AAAG-GAG--QGGYGGLGSQGTSGRGGLGGQGA------GAAAAAAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 154 GQGGYGGLGSQGTSG-----G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AFSGIPGVGPFGCQQPGVPLGYPIKAPKLPGGYGLPYTNGKLPYGVAGAGGKAGYPTGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         189 AGAAAAAAAAA---GGAGQG----GYGGLGSQGTSGR----GGLGGQGAGAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Score 939; DB 1; Length 864;
; Pred. No. 6.8e-36;
32; Mismatches 279; Indels 198;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MISSING (IN CERTAIN ISOFORMS).
MISSING (IN CERTAIN ISOFORMS).
MISSING (IN CERTAIN ISOFORMS).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MISSING (IN CERTAIN ISOFO)
456894BB09E79FD4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Structural protein; Connective tissue; Repeat; Signal;
                                                                                                                                                                                                                                                     -1- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY
                                                                                                                                                                               INTO AN EXTENSIBLE 3D NETWORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M86355; AAA42271.1; JOINED.
EMBL; M86363; AAA42271.1; JOINED.
EMBL; M86364; AAA42271.1; JOINED.
EMBL; M86316; AAA42271.1; JOINED.
EMBL; M86371; AAA42271.1; JOINED.
EMBL; M86376; AAA42271.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    72786 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26.9%;
38.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M60647; AAA42269.1; -. EMBL; J04035; AAA42268.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       M86373; AAA42272.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 316; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M86375; AAA42272.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  864 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                        splicing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARSPLIC
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   353
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
```

```
Conservative
                                                                                                                                                                                                                                                                                                                                                            603 AA;
                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
tes 256; Conserv
    SUBFAMILY
                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            237
  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DoBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                            420
                                                      384
                                                                                 530
                                                                                                                                                                 457
                                                                                                                                                                                        591 GTGLVPGDLGGAGTPAAAKSAAKAAAKAQYRAAAGLGAGVPGLGVGAGVPGFGAGAGGFG 650
                                                                                                                                                                                                                      511
                                                                                                                                                                                                                                                 704
                                                                                                                                                                                                                                                                         705 LGGVPGGVAGGAPAAAAAAAAAAAYGYGLGGAGGLGAGGLG------AGGL 755
                                                                                                                                                                                                                                                                                                                                 626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            'n
GGLGGQGAGAAAAAAAAAGGAGQGYGGLG--SQGTSGPGGYGPGQQTSGRGGLGG--Q 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Harris
                         413 GIVGGPGAVSPAAAAKAAAKAGAKGGVGIPTYGV-GAGGF-PGYGVGAGAGLGGASQ
                                                       -----GSQGTSGRG
                                                                             471 AAAAAAAAKKYGAGGAGTLGGLVPGAVPGALPGAVPGALPGAVPGALPGAVPGTG
                                                                                                                                   531 GVPGAGTPAAAAAAAAAAKAAAKAGQYGLGPGVGGVPGGVGGLPGGVGPGGVTGIGTGP
                                                                                                                                                              421 ----RGGLGGQGAGAAAAAAA-----AAAG-GAGQGGYG------GLGSQGTS
                                                                                                                                                                                                                                              651 AGAGVPGFGAGAVPGSLAASKAAKYGAAGGLGGPGGLGGPGGLGGPGGFG-----GPGG
                                                                                                                                                                                                                                                                                                                              570 GQGGY--GGLGSQGT-SGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                  Holroyd
                                                                                                                                                                                                                   458 GRGGLGGQGAGA----AAAAAAAAAGGAGGGYGGLGSQ-GTSGPGGYGPGQQTSGRGG
                                                                                                           --AGGAGOGY-GGLGSOGTSG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-H37Rv;
MEDLINE-98295987; PubMed-9634230;
Gole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harr.
Goldon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Erown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroy,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
Deciphering the biology of Mycobacterium tuberculosis from the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                   GGQGAGAAAAAAAAAAGGAGQGG----YGG-LGSQGTSGPGGYG 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv1325c precursor.
RV1325C OR MT1367 OR MTCY130.10C.
                                                                                                                                                                                                                                                                                                                                                                                                              812 ARPGFGLSPIYPGGGAGGLGVGGKPPKPYGGALGALGYQGGGCFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   603
                                                      GAGAAAAAAAAGGAGGGGGGGGC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                           GLGGGGAGAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   complete genome sequence.";
Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YD25_MYCTU
Q10637;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bishai W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Whole
                                                                                                                                                                                                                                                                         512
                                                                                                         385
296
                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                 8
                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                           셤
                                                                                                         à
                                                                                                                                                             ò
                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---GAGGAGGR-SGLFEVLAAGGAGGTGGLSVNGGTGGTG---GTGGGGGLFSNGGAGGAG 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          291 GFGVSGSAGGNGGTGGDGGIFTGNGGTGGTGTGTGNQLVGGEGGAGGAGGNAGILFGAG 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      602 AAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGS-QGTSG 660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66 SAQMSAFHAQFVQTFTAGAGAYASAEAAAAPLEGLLNIVNTPTQLLLGRPLIGNGANGA 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               86 GSQGTSGRGGL---GGQGAGA-AAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGA 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142 AAAAAAAAAGGA--GQGGYGGLGSQG-TSGPGG----YGPGQQTSGRGGLGGQGAGAAA 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         363 AAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLG-SQGTSGR 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26 GAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGGYGGL 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            195 AAAAAAGGA--GQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GTSGRGGLGGQGAGAAAAAAAAAAGGAG----QGGYGGLGSQGTSGRGGL---GGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    422 GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              482 GQGGYGGLGSQGTSGPGQQTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     126 PGTGQAGGAGGLLYGNGGAGGSGAPGQAGGPGGAAGLFGNG----GAGGAGGDGPGNGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               306 AAAAAAAAAAGGAGGGYGGL--GSQCTSGPGGYGPGQQ-TSGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     351 GIGGTGGTGLGAPDPGGTGGKGGVGGIG-----GAGALFGPGGAGGTGGFGASSADQMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            542 QGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQCTSGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    494 -ETAGDGGAGGN-----AGLLNGDGGAGGAGGLGIAGDGGNGGKGGK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 GGAGGAGGLLFGSGGAGGPGGVGNTGTGGLGGDG-GAAGLFGAGGIGGAGGPGFNG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTËNTIAL.
HYPOTHETICAL PE-PGRS FAMILY PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 126;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24.9%; Score 869; DB 1; Length 603; 38.0%; Pred. No. 7e-33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RV1325C.

GLY-RICH.

G -> V (IN REF. 2).

G -> D (IN REF. 2).

G -> D (IN REF. 2).

G -> P (IN REF. 2).

G -> P (IN REF. 2).

M; 4F9BCBB2B07AE964 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         35; Mismatches 256;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7e-33;
                                                                                                                                                                                                                                                                                            EMBL; AE007010; AAK45630.1; ALT_INIT.
TIGR; MT1367; -.
TubercuList; Rv1325c; -.
                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000084; PE_region. Pfam; PF00934; PE; 1. ProDom; PD001223; PE_region; 1. Hypothetical protein; Repeat; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              49575 MW;
                                                                                                                                                                                                                                                              EMBL; Z73902; CAA98089.1;
```

```
259
                                                                                                                                                                   415
                                                                                                                                                                                                                                                                                                                                                                                                        420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bos
                                                              g
                                                                                                 ò
                                                                                                                                  g
                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
 원
                                ò
                                                                                                                                                                   ð
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@labs.sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL). SUBCELLULAR LOCATION: Cell wall (Potential). DOMAIN: THE REPEATED DOMAINS OF THE PROTEIN FORM A BETA-PLATED
                                                                                                                                                                                             01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
01-JUL-1984 (Rel. 30, Last annotation update)
01-OCT-1994 (Rel. 30, Last annotation update)
01-OCT-1994 (Rel. 30, Last annotation protein 1.8 precursor (GRP 1.8).
Phaseolus vulgaris (Kidney bean) (French bean).
Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---AGMVGNGGDGGAGGASVVANGGVGGSGGN-----ATLIGNGGNGGNGGNGGAGGAGG 586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLG-GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGA 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GOGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP- 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               235 AAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGGGGYGGLGSQGTSG 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GOGGGVGYGGDQGAGYGGGGGSGGG----GGVAYGGGGERGGYG--GGQGGGAGGGYGAG 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GQQTSGRGCLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAA 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            33 GYGLGHGTGGGYGGAAGSYGGGGGGGGGGG-GYAGEHGVVGYGGGGS-----GG 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: THE N-TERMINAL SIGNAL SEQUENCE OF ABOUT 30 AA SHOWS ABOUT 60% HOMOLOGY TO THAT OF THE GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN GRP 1.0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24.4%; Score 853.5; DB 1; Length 465; 40.8%; Pred. No. 2.9e-32; ive 15; Mismatches 163; Indels 184;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
GLYCINE-RICH CELL WALL STRUCTURAL
PROTEIN 1.8.
                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-cv. Tendergreen;
MEDLINE-89091109; PubWed-3208742;
Keller B., Sauer N., Lamb C.J.;
'Glycine-rich cell wall proteins in bean: gene structure and association of the protein with the vascular system.";
EMBO J. 7:3625-3633(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    _
=
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLY-RICH.

8 X 22 AA TANDEM REPEATS.

B5C4A9B983B43607 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; S01820; S01820.
Cell wall; Structural protein; Repeat; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    =
=
-
-
                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         36683 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X13596; CAA31932.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SHEET CONFIGURATION.
                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          465
                                                                                 587 AGG-----TAGL 593
                                                661 PGGYGPGQQTSGI 673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    =
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           465 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=3885;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          33
                                                                                                                                                               GRP2_PHAVU
P10496;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Sım
Matches 249;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                  RESULT 12
GRP2_PHAVU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
               535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              175
               g
                                                ð
                                                                               g
                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
Rosenbloom J.; "Structure of the bovine elastin gene and S1 nuclease analysis of alternative splicing of elastin mRNA in the bovine nuchal ligament."; Blochemistry 28:2365-2370(1989).
                                                   414
                                                                                                                                                        474
                                                                                                                                                                                     289 AGGEHGGGAGGGOG------GGAG----GGYGAGGEHGGGGGGGGGGGGGYA 332
                                                                                                                                                                                                                                                                                                              419
                                                                                                                                                                                                                                                                                                                                             654
                                                                                                                                                                                                                                                                                594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-87194772; PubMed-3032943;
Raju K., Anwar R.A.;
"Primary structures of bovine elastin a, b, and c deduced from the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Cetartiodactyla, Ruminantia, Pecora, Bovoldea,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-89274159; PubMed-2543440;
Yeh H., Anderson N., Ornstein-Goldstein N., Bashir M.M.,
Rosenbloom J.C., Abrams W.R., Indik Z., Yoon K., Parks W., Mecham
                                                                                                                       ------GGAG----GGYGAGGEHGGGAGGGQG--------GGAG----GGYG
                                                                                                                                                                                                                                     ----GGAGGGGGG-----GDHGGGYGGQQAG-----GGAG-GGYGGGGEHGGGG
                              295 RGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAG
                                                                                          355 AAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLG
                                                                                                                                                        384 GYGG------GQGGGGGYGAG------GDHGAAGYGG------GEGGGGGSGGG
                                                                                                                                                                                                                                                                                                                                           475 AAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAAGGAGG
                                                                                                                                                                                                                                                                                535 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ELS_BOVIN STANDARD; PRT; 747 AA. P04985; P04987; 029421; 13-406-1987 (Rel. 05, Created) 13-406-1987 (Rel. 05, Last sequence update) 16-0CT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sequences of cDNA clones.";
J. Biol. Chem. 262:5755-5762(1987),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INTO AN EXTENSIBLE 3D NETWORK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Elastin precursor (Tropoelastin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1-27 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE-Nuchal ligament;
MEDLINE-89274159; PubMec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bovidae; Bovinae; Bos
                                                                                                                                                                                                                                                                                                                                                                                                         655 SQGTSGPGGYG 665
                                                                                                                                                                                                                                                                                                                                                                                                                                        GIG----GGHG 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             taurus (Bovine).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=9913;
```

```
183 GAGVKPKAQVGAGAFAGIPGVGPFGGQQP-----GLPLGYPIKAPKLPAGYGLPYKTG 235
                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. MEDLINE-84270667; PubMed-6087149;
                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EBNA-1 nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                 705 PIGGGAGGLGVG 716
                                                                                                                                                                                                                                                                                                                                                             658 T--SGPGGYGPG 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=10377;
                                                                                                                                                                                                                                                                                                                                                                                                                                     EBN1_EBV
P03211;
                                                                                                         347
                                                                                                                                                                                            456
                                                                                                                                                                                                                                    508
                         240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BKRF1
                         à
                                           g
                                                                ò
                                                                                   g
                                                                                                         à
                                                                                                                             셤
                                                                                                                                                  ò
                                                                                                                                                                   පු
                                                                                                                                                                                           à
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                          d
                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GGQGAGAAAAAAAAAG-GAG--QGGYGGLGSQGTSGPGGY--GPGQQTSGRGGLGG--Q 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGLGVSTGAVVPQLGAGVGAGVKPGKVPGVGLPGVY-PGGVLPGAGARFPGIGVLPGVPT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLG---SQGTSGRGGLGGQGAGAAAAAAAA 74
SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS. ALTERNATIVE PRODUCTS: 3 ISOPONMS: A (SHOWN HERE), B AND C; ARE PRODUCED BY ALTERNATIVE SPLICING. PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22.2%; Score 775.5; DB 1; Length 747; 39.1%; Pred. No. 1.1e-28; 1ve 27; Mismatches 295; Indels 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 AND 3).
                                                                                                                                                                                                                     InterPro; IPR003979; tropoelastin.
PRINTS; PR01500; TROPOELASTIN.
Structural protein; Connective tissue; Repeat; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             633C03E411643D83 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                OXIDATIVE DEAMINATION
MISSING (IN ELASTIN B).
                                                                                                                                                                                                                                                                                                                                                                                                   DEAMINATION.
DEAMINATION.
DEAMINATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 AND
                                                                                                                                                                                                                                                                                                                              DEAMINATION
DEAMINATION
                                                                                                                                                                                                                                                                                                                                                            DEAMINATION
                                                                                                                                                                                                                                                                                                                                                                        DEAMINATION
                                                                                                                                                                                                                                                                                                                                                  DEAMINATION
                                                                                                                                                                                                                                                                                                                                                                                DEAMINATION
                                                                                                                                                                                                                                                                                                                                                                                           DEAMINATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MRS -> MAG (IN REF.
E -> G (IN REF. 2 AN
                                                                                                                                                                                                                                                                                                                   OXIDATIVE IOXIDATIVE I
                                                                                                                                                                                                                                                                                    OXIDATIVE I
                                                                                                                                  EMBL; J02717; AAA30503.1; --
EMBL; K03505; AAA30505.1; --
EMBL; K03506; AAA30506.1; --
EMBL; J02855; AAA30776.1; --
EMBL; M58652; AAA03519.2; --
PIR; A26728; A26728.
PIR; C26728; C26728.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64229 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                      Alternative splicing.
SIGNAL 1 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12
747 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                  271
273
327
327
400
404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOD_RES
MOD_RES
MOD_RES
MOD_RES
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                  MOD_RES
MOD_RES
MOD_RES
MOD_RES
MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match
                                                                                                                                                                                                                                                                                                                                                                    MOD_RES
MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                           MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                            MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local
Matches 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   188
 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
```

```
Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J., Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C., Tuffnell P.S., Barrell B.G., "DNA sequence and expression of the B95-8 Epstein-Barr virus genome."; Nature 310:207-211(1984).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     543
                                                                                                                                                                                                         296 GAPGAIPGIGGIAGVGAPDAAAAAAAAAAKFGAAGGLPGVGVPGVGVPGVGVPGVGVPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBCELLULAR LOCATION.
MEDLINE-9026473; PubMed-2161150;
MEDLINE-9026473; PubMed-2161150;
MEDLINE-9026473; Sample C., Kleff E.;
"Subnuclear localization and phosphorylation of Epstein-Barr virus latent infection nuclear proteins.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1-26 FROM N.A.
MEDLINE-86259739; PubMed-3460083;
Sample J., Hummel M., Braun D., Birkenbach M., Kieff E.;
Sample J., Hummel M., Braun D., Birkenbach M., Kieff E.;
Processof MRNAs encoding Epstein-Barr virus nuclear
proteins: a probable transcriptional initiation site.";
Proc. Natl. Acad. Sci. U.S.A. 83:5096-5100(1986).
----GAGQGGY-GGLGSQCTSGRGGLGGQGAGAA-AAAAAAAGGAG---QGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                     356 GVGVPGVGVP-----GVGVPGVGVPGVGVPGV-GVPGVGALSPAATAKAAAKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          236 KLPYGFGPGGVAGSAGKAGYPTGTGVGPQAAAAAAAAAKLGAGGAGVLPGVGVGCPGIP
                                                                                                                                    GTSGR----GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRG
                                                                                                                                                                                                                                                                               GLGGQGAGAAAAAAAAAAGGAGQGGYG--GLGSQGTSGRGGLGGQGAGAAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        166 GAIPGLPGVGGVPGVGIPAAAAKAAAKAAAKGFGLGPGVGVAPGVGVVPGVGVVPGVGVAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 602 AAAGGAGQGGYGGLGSQGTSGRGGLGGQ----GAGAAAAAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                          405 AGQGGYG--GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----GRGGLGGQGAGAAAAAAA------AAAG------GAGQGGYG-GLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --TSGRGGLGG-QGAGAAAAAAAAAGGAGQGGYG-GLGSQGTSG--PG-GYGPGQQTS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ₹.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Epstein-barr virus (strain B95-8) (Human herpesvirus Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Lymphocryptovirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           641
```

```
g
                                                                                                                                                                       ð
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         221 GLGGGGAGAAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAG 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        341 QTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAA 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              173 GPGQQTSGR------GGLGGQGAGAAAAAAAAAAAGGAGGGGYGGLGSQGTSGRG 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                         SOMEWHAT
                                                                                                                                                                                  FUNCTION: INVOLVED IN LATENT CYCLE, EBNA-1 FUNCTIONS IN THE MAINTENANCE REPLICATION OF EBV EPISOME. TRANSACTIVATING FACTOR FOR THE CRICIAL AND ENHANCER FUNCTIONS OF ORIP.
SUBCELLULAR LOCATION: NUCLEAR. FREE IN THE NUCLEOPLASM, SOMEWHAA SSOCIATED WITH THE CHROMATIN AND HARDLY, IF AT ALL ASSOCIATED WITH THE NUCLEAR MATRIX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nuclear protein; DNA-binding; Transcription requlation; Activator;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GPGTGPGNGLGEKGDTSGPEGSGGSGPQRRGGDNHGRGRGRGRGRGG--GRPGAPGGSGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 GAGQGGYGGLGSQG-TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              401 AAGGAGQGGYGGLGSQGTSGRGGLGGQGAAAAAAAAAGGAGQGGYGGLGSQGTSGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       domain of the Epstein-Barr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21.8%; Score 763; DB 1; Length 641;
35.7%; Pred. No. 3.5e-28;
tive 22; Mismatches 209; Indels 186;
                                                             MEDLINE-96006523; PubMed-7553871;
Bochkarev A., Barwell J.A., Pfuetzner R.A., Furey W.F. Jr.,
Edwards A.M., Frappier L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLY/ALA-RICH.
; 4D161653E16FC341 CRC64;
                                         X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 469-607.
                                                                                                                         "Crystal structure of the DNA-binding cvirus origin-binding protein EBNA 1."; Cell 83:39-46(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56427 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; V01555; CAA24816.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL, M13941; AAA45889.1; -. PIR, A03773; QQBE31. PIR; S33021; S33021.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Interpro; IPR004186; EBNA1.
Pfam; PF02905; EBNA1; 1.
Virology 176:563-574(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 35.7% Matches 232; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PDB; 1VHI; 23-DEC-96.
TRANSFAC; T00211; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE 641 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 521
    NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                           OSSSSGSPPRRPPGGRRPFFHPVGEADYFEYHOEGGPDGEPDVPPGAIEQGPADDPGEGP 446
                                                                                                                                                                                 STGPRGGGDGGRRKKGGWFGKHRGQGGSNPKFENIAEGLRALLARSHVERTTDEGTWVAG 506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Partial characterization of cDNA clones encoding the three distinct pro alpha chains of type I collagen from rainbow trout."; Fisheries Sci. 64:780-786(1998).
-i- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEATING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AMINO-TERMINAL PROPEPTIDE (POTENTIAL).
COLLAGEN ALPHA 2(I) CHAIN.
CARBOXYL-TERMINAL PROPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -1- SUBUNIT: TRIMERS OF ONE ALPHA 2(1) AND TWO ALPHA 1(1) CHAINS.
-1- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LICAMENTS AND BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
NCBI_TaxID=8022;
                                                                                                                               ---AG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATUNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Fibroblast;
Salto M., Kunlsaki N., Hirono I., Aoki T., Ishida M., Urano N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Probom; PD002078; F1b_collagen_C; 1.
SMART; SM00038; COLFI; 1.
Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-21257802; PubMed-11358497;
Salto M., Takenouchi Y., Kunisaki N., Kimura S.;
"Complete primary structure of rainbow trout type I collage consisting of alpha!(I)alpha2(I)alpha3(I) heterotrimers.";
Eur. J. Blochem. 268:2817-2827(2001).
                                                                                                                                                                                                                                                                                                             507 VFVYGGSKTSLYNLRRGTALAIPQCRLTPLSRLPFGMAPGPGPQPGPLR
                                                                                                                                                                                                                                                  QGGYGGLGS-----GGTS------GPGCY--GPGOTSGIR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Collagen alpha 2(1) chain precursor.
                                                                                                                      612 YGGLGSQGTSGRGGLGG----QGAGAAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Interpro; IPR000087; Collagen.
Interpro; IPR000885; F1b_collagen_C.
Pfam; PF01391; Collagen; 17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (FIBRILLAR FORMING COLLAGEN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [2]
SEQUENCE OF 417-1356 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Collagen; Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AB052837; BAB55663.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1096
1356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HYDROXYAPATITE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              م
1097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3lycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CA21_ONCMY
093484;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COLIAŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                  647
                                                                                                                                                                                                                                                                                                                                                                                                                                          CA21_ONCMY
                                                                                                                                                                                    447
                                                           387
                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
```

```
29;
                                                                                                                                                                                         73 AAGGAGQGGYGGL-GSQGTSGRGGLGGQGAGAAAAAAAAAAGGAG-QGGYGGLGSQGTSG 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         586
                                                                                                                                                                                                                   167 LPGMKGHRGYNGLDGRKGESGTAGAKGETGAHGANGSPGPAGSRGLNGERGRAGPAGPAG 226
                                                                                                                                                                                                                                                                              287 PVGPVGNPGNNGINGAKGAAGLPGVAGAPGFPGPRGPPGPGPGPGSTGARGLGGDPGPSG 346
                                                                                                                                                                                                                                                                                                                                                                                                              -----GSQGTSGRGGLGGQGAGAAAA-----AAAAAAGGAGQGG------YGGLG 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TSG-RGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        707 ERGEGGPAGLPGFAGPPGSDGQSGPRGEKGPAGGKGDVGPAGPSGQSGPSGASGPAG 766
                                                                                                                                                 131 RGGLGGQGAGAAAAAAAAAAGGAG-----10GGYGGLGSQGTSGP-GGYG-PGQQ---- 177
                                                                                                                                                                                                                                                                                                                           ----TSGRGGLGGQGAGAAAAAAAAGGAG-QGGYGGLGSQGTSGRGGLGGQGAGAA 230
                                                                                                                                                                                                                                                                                                                                                                                      231 AAAAAAAAGGGG-YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGL-- 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SQGTSG-PGGYGPGQQ----TSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGL-- 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     467 LDGRTGPPGPTGPRGQPGNIGFPGPKGPGGEAGKGGDKGPTGATGLRGGPGADGNNGAPG 526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        514 --GQGAGAAAAAAAAAAAAAAAAAA 564 564 590 566 566 566 566 567 564 564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAGGAGQGGYGGL-GSQGTSGRGGLGGQGAGAAAAAAAAAAAG-------GAGQG 610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               611 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGGGGGGGGGG-GSQGTSG-PGGYG--- 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                           SGRGGLGGQGAGAAAAAAAAAAAGAAQGGY-GGLGSQGTSG-RGGLGGQGAGAAAAAAA 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------GSQGTSGRGGLGGQGAGAAAAAAAAAGGAGGGYGGL-----GSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     527 PAGVVGNTGEKGEQGPAGAPGFQGLPGPAGPAGEAGKAGNQGMPGDQGLPGPAGVKGERG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  477 AAGGAGQGGYGGL-----GSQGTSGP------GGYGPGQQTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            767 PPGGRGDAGPSGLTGFPGAAGRVGGPGPAGIAGPPGSAGPAGKDGPRGLRGDPGPGGPQG
(BY SIMILARITY).
N-LINKED (GLCNAC. . .) (POTENTIAL).
W; 7BB2F1F80DB10C93 CRC64;
                                                           Query Match 21.4%; Score 748.5; DB 1; Length 1356; Best Local Similarity 31.6%; Prcd. No. 2.3e-27; Matches 253; Conservative 44; Mismatches 368; Indels 135;
                                126985 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -PGQ----QTSGIRRPAAKL 680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              887 EPGRLGPAGASGPRGPAGNI 906
                1257
             1257 125
1356 AA;
               CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                          178
                                                                                                                              15
                                                                                                                                                                                                                                                                                                                                                                                                                        347
                                                                                                                                                                                                                                                                                                                                                                                                                                                        288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      418
FT
                                                                                                                                                           g
                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

Search completed: December 18, 2002, 16:18:04 Job time: 20.4528 secs

```
GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2002, 16:13:04; Search time 39.982 Seconds

(without alignments)
3509.535 Million cell updates/sec

Title: US-09-490-291-4

Perfect score: 3493
Sequence: 1 MRGSHHHHHHGSMASGRGCL.......GGYGPGQQTSGIRRPAAKLN 681

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Database : SPTREMBL_21:*

1: SP_archea:*
2: SP_bacteria:*
3: SP_langi:*
4: SP_lungi:*
5: SP_lungi:*
7: SP_lungi:*
7: SP_lungi:*
8: SP_lungi:*
7: SP_lungi:*
17: SP_lungi:*
10: SP_lungi:*
11: SP_lungi:*
11: SP_lungi:*
12: SP_lungi:*
13: SP_lungi:*
14: SP_lungi:*
15: SP_lungi:*
16: SP_lungi:*
17: SP_lungi:*
18: SP_lungi:*
19: SP_lungi:*
11: SP_lungi:*
11: SP_lungi:*
12: SP_lungi:*
13: SP_lungi:*
14: SP_lungi:*
16: SP_lungi:*
17: SP_lungi:*

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

						٠									
Description		046171 nephila cia	Q9biu3 dolomedes t	Q9biu7 argiope tri	Q9biu4 dolomedes t	-	O9bit7 nephila mad		Q964f4 antheraea y	Q8v1y9 mycobacter1	Q8viz0 mycobacteri	053559 mycobacteri	053557 mycobacteri	Q9bivl argiope aur	053775 mycobacteri
SUMMARIES	Q8WSW4	046171	Q9BIU3	Q9BIU7	Q9BIU4	017434	Q9BIT7	076786	Q964F4	Q8VIY9	Q8VIZO	053559	053557	Q9BIV1	053775
DB	ហេដ	n n	'n	'n	S	'n	S	'n	'n	16	16	16	16	Ŋ	16
% Query Watch Length DB	644	544	691	648	854	988	1953	2639	2655	1217	1715	1489	1079	447	1306
& Query Match	57.5	49.4	45.6	44.9	43.8	41.5	41.5	39.7	38.6	36.3	36.2	35.9	35.2	34.3	34.3
Score	2009	1726.5	1592	1568	1531.5	1450	1448	1386.5	1350	1267	1263.5	1254.5	1228.5	1199.5	1197
Result No.	140	1 M	4	2	ø	7	8	6	10	11	12	13	14	15	16

Q8v1z1 mycobacteri O53552 mycobacteri O53215 mycobacteri Q8v123 mycobacteri Q9nhv2 neplila mad	മഥ ഭത	O'053439 mycobacteri 053844 mycobacteri 04359 nephila cla 044358 nephila cla 044358 nephila cla 080x42 mycobacteri 053845 mycobacteri 053415 nephila mad	053684 mycobacteri 053684 mycobacteri 053809 mycobacteri 016987 arancus dia 08vj15 mycobacteri 050415 mycobacteri 053556 mycobacteri 006808 mycobacteri 09esz9 mus musculu
Q8VIZ1 O53552 O53215 Q8VJ23 Q9NHW2	002402 003395 080K17 090HW4 09GUB5	053449 053844 054359 044358 08VKD2 053845 053145	053684 09BIU9 0153809 016987 08WJ15 050415 053556 006808
16 16 16 5	22,000	16 16 16 16 16 16 16 16 16 16 16 16 16 1	110 110 110 110 110 110 110 110 110 110
1384 1381 1665 1507 1884	1538 1408 1468	1327 1853 149 871 882 563 767	837 651 783 775 731 741 810
# # # # # # # # # # # # # # # # # # #		226 231.6 331.9 230.9 29.8	78888888888888888888888888888888888888
1186.5 1182.5 1173 1160.5 1159.5	1157 1157 1152.5 1151 1148	1114.5 11114.5 1099 1083 1080.5 1075 1041.5	1015.5 1015.5 1014 1011 1005 1004 984.5
		9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

ALIGNMENTS

~

```
Beckwitt R., Arcidiacono S., Stote R.; "Evolution of repetitive proteins: spider silks from Nephila clavipes (Terragnathidae) and Araneus bicentenarius (Araneidae)."; Insect Biochem. Mol. Biol. 28:121-130(1998).
109 -AGRGGSGGQGAG----AAAAAAGGAGQGGYGGLGSQG------AGRGGLGGQ
                                                                                                                                                               AGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGGYGGLGSQGTSGRGG
                                                                                                                                                                                                                                          LGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAA
                                                                                                                                                                                                                                                                                                                        249 RGGQGAGAAAAAAV----GAGQGGYGGQG----AGQGGYGGLGSQGAGRGCLGGGGAG--
                                                                                                                                                                                                                                                                                                                                                                                                     AAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GAGAAAAAAAAAAAG------GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                G-AGRGGQG-----AAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAG----AAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     596 AAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGGLGS
                                                            TSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Araneae;
Nephila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Arcidiacono S., Mello C., Kaplan D., Cheley S., Bayley H.;
"Purlification and characterization of recombinant spider silk
parmessed in Bacherichia coli ".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Xu M., Lewis R.V.;
"Structure of a protein superfiber: spider dragline
Proc. Natl. Acad. Sci. U.S.A. 87:7120-7124(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appl. Microbiol. Biotechnol. 49:31-38(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
MEDLINE-90384959; PubMed-2402494;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 449-544 FROM N.A. MEDLINE-98148687; PubMed-9487707;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                spider).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                expressed in Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUN-1998 (TrEMBLrel. 06, 01-JUN-1998 (TrEMBLrel. 06, 01-Dec. 2001 (TrEMBLrel. 19, Spidroin 1 'P---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Spidroin 1 (Fragment).
Nephila clavipes (Orb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [3]
SEQUENCE FROM N.A.
                    AAAAAAGGVGO-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=98318805;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Araneomorphae; E
NCBI_TaxID=6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             046171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            046171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       236
                  63
                                                                                                                                      188
                                                                                                                                                                              151
                                                                                                                                                                                                                       238
                                                                                                                                                                                                                                                          207
                                                                                                                                                                                                                                                                                                  298
                                                                                                                                                                                                                                                                                                                                                                              357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           485
                                                          128
                                                                                                                                                                                                                                                                                                                                                                                                                                                           417
                                                          à
                                                                                                                                                                         g
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HERER REPRETATION OF THE REPRETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           [1]
SEQUENCE FROM N.A.
MEDLINE-94165058; PubMed-8120021;
BECKWILT R., Arcidiacono S.;
"Sequence conservation in the C-terminal region of spider silk
"Sequence statements (Spidroin) from Nephila clavipes (Tetragnathidae) and Araneus
bloentenarius (Araneidae).";
J. Biol. Chem. 269:6661-6663(1994).
                                                                                                                                                                              310
                                                                                                                                                                                                                                                                                                  486
                                                                                                                                                                                                                                                                                                                                                                              545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Beckwitt R., Arcidiacono S., Stote R.;
"Evolution of repetitive proteins: spider silks from Nephila clavipes
(Tetragnathidae) and Araneus bicentenarius (Araneidae).";
Insect Blochem. Mol. Blol. 28:121-130(1998).
                                                          307
                                                                                               254
                                                                                                                                                                                                                                                          354
                                                                                                                                                                                                                                                                                                                                       401
                                                                                                                                                                                                                                                                                                                                                                                                                    444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GSMASGRGGLGGQGA---GAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA 67
487 GGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTS
                                                                                                                                                                                                                                                                                                                                                                                                   --GGLGGQGAGQGAGAAAAAAGGVRQGGYGGLGSQG----AGRGGQGAG----AAAAAAG
                                                       GLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA
                                                                                                                                    AAAAAAGGAGGGYGGLGSQGTSGPCGYGP-GQQTSGRGGLGGQGAGAAAAAAAAGG
                                                                                                                                                                                                                 AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                    546 GRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAA
                                                                                                                                                                                                                                                                                              427 QGAGAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGGGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          44; Indels 147;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           49410 MW; 1EF0CE9269A832E2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Dragline silk protein spidroin 1 (Fragment).
Nephila clavipes (Orb spider).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 57.0%; Score 1991.5; DB 569.2%; Pred. No. 6.6e-100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; U37520; AAC04504.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                --- AAAAAAAGGAGO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         617 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Araneomorphae; E:
NCBI_TaxID=6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUN-1998
01-JUN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 457;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              046172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   046172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      909
                                                       248
                                                                                                                                    308
                                                                                                                                                                         255
                                                                                                                                                                                                                   367
                                                                                                                                                                                                                                                                                                                                     355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
046172
                合
                                                       ð
                                                                                           g
                                                                                                                                  à
                                                                                                                                                                       g
                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

476

535

441 595 484 655 534

416

343

206

237

248

356

297

```
Spider Silk
                                                                                                                                                               NON_TER 1 1
SEQUENCE 691 AA; 56176 MW; CB46EF9F1991F33E CRC64;
    and Convergence of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GAGAAAAAAAAAG-GAGQGGYGGLGSQGTSGPGGYGPG 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last seque:
01-DEC-2001 (TrEMBLrel. 19, Last annot.
Major ampullate spidroin 1 (Fragment).
MASPI.
"Extreme Diversity, Conservation, an Fibroin Sequences.";
Science 291:2603-2605(2001).
EMBL, AF550270; AAX30599.1;
InterPro: IPR002019; Eggshell.
PRINTS; PR00308; ANTIFFEEEI.
PRINTS; PR01228; EGGSHELL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9BIU7
Q9BIU7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9BIU7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GERFFE
    RTT BE DE RTT DE
                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                     27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        280
                                                                                                                                                                                                                                                                                                                                                                                                        GGQGAGAAAAAAAAAGGAGGGYGGCGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAA 194
                                                                                                                                                                                                                                                                                                                                                                                                                                           93 GGQGAGQGAGAAAAAAGGAGGGYGGLGSQG-----AGRGGSGGQGAG---- 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   255 SGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAA 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        315 GCAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGG 374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              375 LGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     403
                                                                                                                                                                                                                                                                                                                        GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL 134
                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                          195 AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              435 AAAAAAGGAGGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------GGLGGQGAG----AAAAAAGGAGQGGLGGQGA-------GQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           352 AGAAAAAAAAAGGVRQGGYGGLGSQG----AGRGGQGAG----AAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dolomedes tenebrosus.
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Lycosoidea; Pisauridae; Dolomedes.
                                                                                                                                                             DB 5; Length 544;
                                                                                                                                                                                                     23; Indels
                                       the EMBL/GenBank/DDBJ databases
                                                                                                                     CC611B3551945615 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE-21179804; Pubmed-11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              49.4%; Score 1726.5; DB 63.3%; Pred. No. 9.3e-86; iive 7; Mismatches 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                   544 AA; 44107 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9BIU3;
01-JUN-2001 (TrEMBLrel. 17,
01-JUN-2001 (TrEMBLrel. 17,
01-DEC-2001 (TrEMBLrel. 19,
                                   Submitted (JAN-1995) to th
EMBL; U20329; AAC38957.1;
HSSP; P10969; 1WGT.
                                                                                                                                                                              Best Local Similarity 63.3
Matches 406; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fibroin 2 (Fragment)
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=156846,
                                                                                                NON_TER
SEQUENCE
                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9BIU3
                                                                                                                                                                                                                                                                                                                                                                                                        135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               404
RA RE DR ET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ο
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 硆
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
```

```
40,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   297 GLG-GQGAGAAAAAAAAAG-GAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAG 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             296 AGAAAAAAGGAGAGGGGGG---QGGYGQGG---GAGAAAAAAAASGGSGGGGGYGGQ 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -GLGSQGTSGRGGLG--GQGAGAAAAAAAAAG-GAGQGGYGGLGSQGTSGRGGLG--GQ 629
                                                                                                                                                                      70 AAAAAGGAGGGGYGGLGSQGTSGRGGLGGQ----GAGAAAAAAAAAG-GAGQGGYGGLG 124
                                                 Gaps
                                                                                                                               61
                                                                                                                             2 GQGSGAGAAAAAAAAGGAGQSGSGPYGASYLSSTTYTTSSQGAGGCVGGYGQGSGTGSAA
                                                                                                                                                                                           AGQGGYGGLG------SQGTSGRGGLGGQGAGA-AAAAAAAAAGGAGQGGYG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22 GOGAGAAAAAAAAAGGAGQGGYGGLG------SOGTSGRGGLGGOGAGAAAA
                                                                                                                                                                                                                                                      125 SQCTSGRGGLG--GQGAGAAAAAAAAG-GAGQGGYGGLGSQGTSGPGGYGPGQQTSGR
                                                                                                                                                                                                                                                                             G-GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAG-GAGQGGYGGLGSQGTSGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       355 AAAAAAAAAAGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAG-GAGQGGYGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       414 GSQGTSGRGGLG-GQGAGAAAAAAAAAG-GAGQGGYGGLGSQGTSGRGGLG-GQGAGAA
Query Match 45.6%; Score 1592; DB 5; Length 691;
Best Local Similarity 57.8%; Pred. No. 1.8e-78;
Matches 404; Conservative 34; Mismatches 83; Indels 178;
```

4

us-09-490-291-4.rspt

```
01-JUN-2001 (TrEMBLrel. 17, 01-JUN-2001 (TrEMBLrel. 17, 01-DEC-2001 (TrEMBLrel. 19,
                                      Dolomedes tenebrosus,
                            Fibroin 1 (Fragment)
                                                                                                                                                                 854 8:
                                                                                                                                                         NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                 246
                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                               149
                                                                                                                                                                                                                                                                                                                                                                                                                 306
                                                                                                                                                                                                                                                                                                                                                                                                                                  301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      482
    g
                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                        30;
                                                                                                                                                                                                                                                                                                                              246
                                                                                                                                                                                                                                                                                                                                                       147 AAAAGG-AGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQ 205
                                                                                                                                                                                                                                                                                             186
                                                                                                                                                                                                                                                                                                                                            GGYGGLGSQGTSGPGQTSGRGGLGGQGAGAAAAAAAGGAGQGGx--GGLGS 377
                                                                                                                                                                                                                                                                                                                                                                                        493
                                                                                                                                                                                                                                                                                                                                                                                                                                 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                  551
                                                                                                                                                                                                                                                                                                                                                                                                                                                        571
                                                                                                                                                                                                         Gaps
                                                                                                                                                                                         46
                                                                                                                                                                                                                           82
                                                                                                                                                                       GQGAGAAAAAAAAAAGG-AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AAAAAAAAGGGGGGGGYDGLGSQG----AGQGGYGQGGAAAAAAAGGA------
                                                                                                                                                                                                                                                                           GGY--GGLGSQGTSGRGGLGGQGAGAAAAAAAAAGG-AGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                             OGAGAAAAAAAAAAGGAGQGGYGGLGS-QG-TSGRGGLGGQGAGAAAAAAAAAAGG-AGQ
                                                                                                                                                                                                                                                                                                                                                                               QGTSGRGGLGGQGAGAAAAAAAAAGG-AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                AAAAGGAGQGGY - - GGLGSQGTSGRGLGGQGAGAAAAAAAAAAAAGG - AGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                TSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQG--GYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GYGGLGSQGTSGRGGLGG-----QGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGG
                                                   MEDLINE-21179804; PubMed-11283372; ° Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.; Extreme Diversity, Conservation, and Convergence of Spider Silk Fibroin Sequences."
                                                                                                                                      44.9%; Score 1568; DB 5; Length 648; 58.5%; Pred. No. 3.2e-77; 1ve 35; Mismatches 128; Indels 112;
Argiope trifasciata.
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Ara
Araneomorphae; Entelegynae; Araneoidea; Araneidae; Argiope.
                                                                                                                      648 AA; 52887 MW; ACDB67067B2302CB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               854
                                                                           Fibroin Sequences.";
Science 291:2603-2605(2001).
EMBL; AF350266; AAK30595.1;
HSSP; P10969; INGT.
                                                                                                                                              Best_Local Similarity 58.59
Matches 388; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                          NCBI_TaxID=156845;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   664 YGP 666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11
572 GGP 574
                                                                                                                      SEQUENCE
                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09BIU4
09BIU4;
                                                                                                                                                                                                        87
                                                                                                                                                                                                                                                                                                                                                             247
                                                                                                                                                                                                                                                          83
                                                                                                                                                                                                                                                                            206
                                                                                                                                                                                                                                                                                                                                                                              378
                                                                                                                                                                                                                                                                                                                                                                                                              437
                                                                                                                                                                                                                                                                                                                                                                                                                                                494
                                                                                                                                                                                                                                                                                                             263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     O9BIU4
ID OS
 à
                                                                                                                                                                                       q
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
```

```
186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 GSHHHHHHGSMASGRGGLGGQGAGAAAAAAAAAG-GAGQGGYGGLGSQGTSGRGGLG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83 GOGAGAAAATAAGGAGSGOGDYGDQGGLGGYGOGSGAGSATAPAAGGSGFGQGGFGNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAGGAG---QGGY---GGLGSQGT------SGPGGYGPGQQTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            195 AAGGGDGYEQGGYGNQGGLGSFGQGAGAAAAASAGGAGSGRGGYG-----DQCGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               249 YGQGAGAAAAAAGGDGYGQGYYGDQG--GRGGY-GQGSGAGSATAAAAGGAGFGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Y - - GGLGSQGTSGRGGLG - - GQGAGAAAAAAAAAG - CAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---SGPGGYGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GQQTSGRGGLG--GQGAGAAAAAAAAAG-GAGQGGYG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---GLGSQGTSGRGGLG--GQGAGAAAAAAAAG-GAGQGGY---GGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  187 QGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAA CAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGDGYGQGGYGNQGGLGSFGQGAGAAAAASAGGAGSGRGGYGQGGYGNQ----GGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --GQGGYG---GLGSQG-----TSGPGGYGPGQQTSG-RGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAGGA - - - GQGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAGG - AGQGGYGGLGSQGT
                                                                                                    Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Lycosoidea; Pisauridae; Dolomedes.
NCBI_TaxID=156846;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 5; Length 854;
                                                                                                                                                                                                                                                                                                                                                                                                        854
71881 MW; 5C52F1B0AE31C6C2 CRC64;
                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE-21179804; PubMed-11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Fibroin Sequences, ";
Science 291:2603-265(2001).
Science 291:2603-265(2001).
EMBL; AP$50269; AAK30598.1; -.
InterPro: IPR000217; Tubulin.
PROSITE: PS00227; TUBULIN: UNKNOWN_1.
                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match 43.8%; Score 1531.5; DB 5; Best Local Similarity 50.7%; Pred. No. 3.6e-75; Matches 412; Conservative 46; Mismatches 143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QGAGAAAAAAAAAGGA---GQGGY---GGLGSQGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GQGAGAAAAAAAAAG-GAGQGGYG-------
Created)
```

48;

```
-GLGS-----QGTSGRGGLG-GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      546 GGYGPGQGPGAAAAAAGPGGYGPGQQGTGAAAAAAGSGAGGYG-----PGQQG 598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GPGQQGPGAAAAASAGRGPGGYGPGQQGPGGPGAAAAAAGPGGYGPGQQGPGAAAAAAG 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAAAAAGRGPGGYGPGQQGPGGSGAAAAAAGRGPGGYGPGQQGPGGPGAAAAA 958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---GLGGGGAG 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            41.5%; Score 1448; DB 5; Length 1953;
38.3%; Pred. No. 2e-70;
Live 25; Mismatches 173; Indels 530; Gaps
                                             -AAAAAAGGGGGGGGGGGG-----TSGP
GAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84 GLGSQG------TSGRG-----GLGGQGAGAAA------AAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        115 AGQGGYG------GLGSQGTSGRG------GLGGQGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               779 GAAAAAAGSGPGGYGPGQGPGRSGAAAAAAAAGRGPGGYGPGQQGPGCPGAAAAAGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGYGPGQQGPGAAAAAAAAAAGRGPGGYGPGGSGAAAAAAAGRGPGGYGPGQQGPGGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAAAAAGGAGQGGYG----GLGSQGTSGRG--------GLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
"Extreme Diversity, Conservation, and Convergence of Spider Silk Fibroin Sequences.";
Science 291:2663-2665(2001).
EMBL; AF350276; AAK30605.1; -.
                                                                                                                                                                                                                                                                                                                                                  Nephila madagascariensis.
Eukaryota, Metazoa, Arthropoda, Chelicerata, Arachnida, Araneae,
Araneomorphae, Entelegynae, Araneoidea, Tetragnathidae, Nephila.
NCBI_TaxID=115969;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1953 1953 159383 MW; C82B4DACD043C3BD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGYGPGQQ-----TSGRG-----GLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                   01-JUN-2001 (TrEMBLrel. 17, Last sequence update) 01-JUN-2001 (TrEMBLrel. 17, Last annotation update) Major ampullate spidroin 2-like protein (Fragment).
                                                                                                                    769 AGAGAAAAGAGSGGAGGYGRGAGAGAAAGAGAAGAGAGAGG
                                                                                                                                                   QGAGAAAAAAAAAGGAGGGYG-GLGSQGTSGPGGYGPG 667
                                                                                                                                                                             818 YGAGAGAAAAAGA--GAGRGGYGGYGGGGGGGGGGG 855
                                                                                                                                                                                                                                                          PRT; 1953 AA.
                                                                                                                                                                                                                                                                                        (TrEMBLrel. 17, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-21179804; PubMed-11283372;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 41.5%, Best Local Similarity 38.3%, Matches 451; Conservative
                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                          09BIT7
                                                                                        276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              170
                                                                                                                                                                                                                             RESULT 8
                                                                                                                                                                                                                                                           à
                                                       g
                                                                                        ò
                                                                                                                    a
                                                                                                                                                   ð
                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25;
                                                                                                                                                                                                                                                                                                                                          MEDLINE-98200471; PubMed-9541398;
Colgin M.A., Lewis R.V.;
"Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like 'spacer regions'.";
Protein Sci. 7:667-672(1998).
EMEL: AF027735; AAC14589.1; -.
InterPro; IPR000817; Prion.
PRINTS: PR00341; PRION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              220 GGYGAG---AGAVAAAGAAGGYG----RGAGGYGGQGGYGAGAGAGAAAA--GAGAG 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGGGGYG-----GLGSQGTSGRGGLG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        196 AAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGGYGGLGSQGTS 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GRGGLGGQGAGAAAA-AAAAAGGAGQGGYGGLG-----SQGTSGRGGLGGQGAGAA 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   423 GRGAGAGAGAGAAAARAGAGAGGAGYGGQGGYGAGAGAGAAAAAAGAGAGGAGGYGRGAG 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      651
                            SGRGGLGGQGAGAAAAAAAAAAG-GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21 GGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              364 AGGAGOGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAAAGGAGOGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   462 LGGQGAGAAAAAAAAAAGGAGQGGYG---GLGSQGTSGPGGYGPGQQTSGRGGLGGQGA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       413 LGSQG------TSGRGGLGGQGAGAAAAAAAAAAAGGGGGYGGLGSQGTSGRGG
                                                                                                                                                                                                       01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-JAN-1998 (TrEMBLrel. 17, Last annotation update)
Minor ampullate silk protein MiSpl (Fragment).
Nephila clavipes (Orb spider).
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Score 1450; DB 5; Length 988;
; Pred. No. 9.3e-71;
17; Mismatches 212; Indels 117;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE 988 AA; 79082 MW; 461E03DF53F7085D CRC64;
                                                                        AGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGI 673
                                                                                                    AGGAG-GGQGGYGQGGYGQGTGAGGASSAGL 775
                                                                                                                                                                                988
                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41.5%;
50.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 354; Conservative
                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local 9
                                        691
                                                                        642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 307
                                                                                                                                                RESULT 7
                                                                                                                                                                             g
                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

51;

us-09-490-291-4.rspt

```
34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAAASGARGSGGYGGYGSDSAAAAAAAAAAAAGSGAGGYGGGYGWGDGGYGSDSAAAA 1081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        849
                                                            135 RGGGFYETHNSYSSYGSGSSSAAAGSGAGGVGGGYGSDSAAAAAAAAAAAAAAGGGGGGGG 194
                                                                                            164
                                                                                                                                                                                       315 AGGAGGGYGWGDGGYGSDSAAAAAAAAAASSGAGGRGDGGYGSGGSSAAAAAAAAA 374
                                                                                                                                                                                                             201
                                                                                                                                                                                                                                550
                                                                                                                                                                                                                                                                                                                                     T------SGRGGLGGGGAGAAAAAAAGGGG--QGGY----GGLGSQGT--- 330
                                                                                                                                                                                                                                                                                                                                                          610
                                                                                                                                                                                                                                                                                                                                                                               378
                                                                                                                                                                                                                                                                                                                                                                                           410
                                                                                                                                                                                                                                                                                                                                                                                                                                          670 SAAAAAAAAAAAAGSGAGGVGGCYGWGDGGYGSDSAAAAAAAAAAAAGSGAGGRGDGGY 729
                                                                                                                                                                                                                                                                                                                                                                                                                                                               411 GGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAAGGAG-QGGY----GGLG 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SQGT-----SGRGGLGG------QGAGAAAAAAAAAAAAAAAAAAA 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 39.7%; Score 1386.5; DB 5; Length 2639;
Best Local Similarity 40.4%; Pred. No. 5e-67;
Matches 402; Conservative 23; Mismatches 239; Indels 330; Gaps
                                       2 RGSHHHHHHGSMA-----SGRGGLGG----QGAGAAAAAAAAAAAAGGAGQGGYG 45
                                                                                  86
                                                                                                                                                                                                                                                                                                                                                        551 AAAAAAAASAARRAGHDSAAGSAAAAAAAAAAAAAAAAAAAAAGGSGGGYGWGDGGYGSDSAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLGSQ-------GTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLG
                                                                                                                          87 SQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTS---------
                                                                                                                                                                    ------GRGGLGGQGAGAAAAAAAA----AGGAGQGGYGGLGSQ------
                                                                                                                                                                                                            ------GTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAA
                                                                                                                                                                                                                                                                                                        -----SGPGGYGPGQOTSGRGGLGGQGAGAAAAAAAAA--GGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGLGSQCT-----SGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAAAAGGAG--QGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----LGSQGTSGRGGLGGQGAGAAAAAAAAGGAGGGGY-----GGLGSQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                      GT---------SGRGGLGG------QGAGAAAAAAAA---AGGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------GTSGPGGYGPGQQTSGIRRPAA 678
                                                                                46
                                                                                                                                                                                                                                                     202
                                                                                                                                                                                                                                                                                            236
                                                                                                                                                                                                                                                                                                                                      292
                                                                                                                                                                                                                                                                                                                                                                                                                     379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1022
                                                                                                                                                                   130
                                                                                                                                                                                                            165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      696
                                                                                                                                                                                                                                                                                                                                                                              331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    657
                                                         qq
                                                                                á
                                                                                                    g
                                                                                                                          ð
                                                                                                                                           g
                                                                                                                                                                  å
                                                                                                                                                                                  g
                                                                                                                                                                                                     ò
                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                        οŽ
                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
```

_

```
40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1368 AAAAAAAGSGAGGAGGGYGWGDGGYGSDSAAAAAAAAAAAAAGGGGGGGGGGGGSSSA 1427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SAAAAAAAAAAAAASGAGGRGDGGYGSGSSAAAAAAAAAAAAAAAARRAGYDRAAGSAAAAAAA 1606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GTSGRGGLGG------QGAG 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----GLGSQGTS---GRGGLGGQGAGA 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    141 AAAAAAAAAAAGAGGGYGGLGSQGT-----SGPGGYGPGQQTSGRGGLGGQ 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             401
                                                                                                     Antheraea yamamai.
Bukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Saturniidae; Saturniinae; Saturniini; Antheraea.
NCBI_TaxID=7121;
                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                      90
                                                                                                                                                                                      SEQUENCE FROM N.A.
TRANSPOSON=MARINER LIKE ELEMENT;
Lee J.-S., Hwang J.-S., Kang S.-W., Suh D.-S., Jin B.-R., Kim Y.-S.,
Lee K.-S., Goo T.-W., Yun E.-Y., Kim K.Y.;
"Complete nucleotide sequence of fibroin gene cloned from Antheraea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1308 AAAAGSGAGGVGGGYGWGDGGYGSDSAAAAAAAAAAAAAAAGGGGGGGGGGGGGGGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----SGRGGLGG-----QGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           188 GAGAAAAAAAAA----AGGAGQGGYG-----G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            211 LGSQGTSGRGGLGG------QGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----SGRGGLGG-----QGAGAAAAAAA---AGGAGQGGYG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LGGQGAGAAAAAAAA--AGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAA----
                                                                                                                                                                                                                                                                                                                                                         448;
                                                                                                                                                                                                                                                                                                                                                                                  GSMASGRG----GLGGQGAGAAAAAAAAAAGGAG--QGGYG-GLGSQGT------
                                                                                                                                                                                                                                                                                                                           38.6%; Score 1350; DB 5; Length 2655; 37.5%; Pred. No. 4.5e-65; Live 22; Mismatches 217; Indels 448
                                                                                                                                                                                                                                                                     Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AF325500; AAK83145.1; -
SEQUENCE 2655 AA; 218672 MW; 89BB810EF9CC4CA6 CRC64;
                                                                 Last sequence update)
Last annotation update)
                          PRT; 2655 AA.
                                                     Created)
                                                  (TrEMBLrel. 19, C
(TrEMBLrel. 19, I
(TrEMBLrel. 19, I
                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 57.59
Matches 412; Conservative
                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                           Similarity
                                                  01-DEC-2001 (
01-DEC-2001 (
01-DEC-2001 (
                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                          Fibroin.
                        Q964F4
Q964F4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  255
                                                                                                                                                                                                                                                                                                                                                                                  11
                                                                                                                                                                                                                                                                                                                                                                                                                                      53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1787
RESULT 10
            Q964F4
                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
STRAIN=CDC 1551 / OSHKOSH;
STRAIN=CDC 1551 / OSHKOSH;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                2074 AANAAARRAGHDRAAGSAAAAAAAAAAAAASGAGRGGGGGGGGSSDSAAAAAAAA 2133
                                                                                                                                                                                                                                                                                                                                                                                   2194 AAAAAAGSGAGGVGGGYGWGDGGYGSDSAAAAAAAAAAAASGAGGSGGYGGYGSDSAAAAA 2253
                               AGGRGDGGYGSGSSAAAAAAAAAAARRAGHDRAAGSAAAAAAAAAAAAAASGAGGSGGGY 1957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    260 GHGGVGGDGGWLAPGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGLG317
------4LGSQGTS- 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GRGGLGGQG----AGAAAAAAAAAAAAGGAGQ-----GGYGGLGSQGTSGRGGLGGQGAGA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                              AAGGAGQGGY-----GGLGSQ-----GTSGPGGYGPGQ-----
                                                                                                                                                                                                                                                          ----GLGSQGTS-----GRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT----
                                                                 --GRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTS-GRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                       ----SGRGGLGG-----QGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 180;
                                                                                                                                                                                            -----QTSGRGGLGGQGAGAAAAAAAAGGAGQ--GGYG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 1217;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: AE007164; AAK47974.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ProDom; PD001223; PE_region; 1.
PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_3.
SEQUENCE 1217 AA; 94463 MW; 2F5BDCA5A30A056E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 36.3%; Score 1267; DB 16; Best Local Similarity 40.1%; Pred. No. 6.9e-61; Matches 328; Conservative 38; Mismatches 272;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 1217 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000084; PE_region.
InterPro; IPR002173; PfkB.
Pfam; PF00934; PE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --- GYGSGSSAAAAAAAA 2329
                                                                                                                                                                                                                                                                                                                                                                                                                                                    GPGGYGPGQQTSGIRRPAA 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    O1-UN-2002 (TENBLE) 20
01-UN-2002 (TENBLE) 21
PE_PGRS family protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAR-2002 (TrEMBLrel. 01-MAR-2002 (TrEMBLrel.
 AGGAGGGGYG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bishai W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08VIY9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8VIY9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2314
402
                                                               420
                                                                                                                             477
                                                                                                                                                                                          505
                                                                                                                                                                                                                                                                                                                       583
                                                                                                                                                                                                                                                                                                                                                                                                                                                    099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                          qq
                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                             ð
                                                                                                                                                       g
                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                             ò
```

```
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [1] ...
SEQUENCE FROM N.A.
STRAIN-CDC 1551 / OSHKOSH;
STRAIN-CDC 1551 / OSHKOSH;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Peterson J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      378 SGAAGNAGVGGAGGGG--GDGGAGGAGADADQPGATGGTGFAGGAGGAGGAGGSSGAGGT 435
                                                                                                                                                                                                                                                                        NGSGGAGGTNGSGCAGGTGGGGAGGAGGAGADNPTGIGGTGGDGGTGGAGGAGGAGGAG 495
                                                                                                                                                                                                                                                                                                                                                                                  615
                                                                                                                      -----GAAAAAAAAGGA 153
                                                                                                                                                                                                                                                                                                                                                    GQGGYGG-LGSQGTSGRGGLGGQGAGAAAAAA------AAGGAGQGGYGGL 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----GSQGTSGRGGLGGQGAGAAAAAA-----AAAGGAGQGGYGGLGSQGT---SGR 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGLGGQGAGAAAAAAAAAGGAGGGYGG-LGSQGTSGPGGYGPGQQTSGRGGLGGQGA-353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GAAAAAAAAAAGG-------380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            673 ADOPGATGGTGFAGGAGGAGGAGGNSGAGGTNGSGGAGGTGGQGGAGGAGGAGADNP 732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  33 TGIGGTGGDGGTGGAGGAGGAGGAGGAGGTGGMIGTTGNAGVGGAGGGGGGGGGAGAGA 792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         318 LGAGGOGGLGGAGGOGGMGGAGGAGADNPTGIGGTGGDGGTGGSAGEGGAGGAAGQLFSA 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      381 SGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG-LGSQGTSGRGGLGGQ----GAGAAAA 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          436 A------AAAAGGA-GQGGYGGL----GSQGTSGRGGLGGQGAGAAAAAAAA 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           535 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAG 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAAAAAAAGGA-----GQGGYGGLGSQGTSG------RGGLGGQG-----AGAA 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mycobacterium tuberculosis.
Bacteria, Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAAAAAAGGA-GQGGYGGLGSQ-----GTSGRGGLGGQ----GAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----GQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          480 GAGQGGYGGLGSQGTSGPGG-YGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGA----GQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1028 GTNGSGAGGAGGAGGAGISFSNGSNGGTGGTGGVG 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               635 AAAAAAAGGAGQGGYGGLG----SQGTSGPGGYG 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 1715 AA
                                                                                                                   AGGACQGGYGGLGSQGTSGRGGLGGQGA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 O1-JUN-2002 (TrEMBLE-1. 20, 01-JUN-2002 (TrEMBLE-1. 21, PE_PGRS family protein. MTS612.1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAR-2002 (TrEMBLrel. 20, 01-MAR-2002 (TrEMBLrel. 20, 01-JUN-2002 (TrEMBLrel. 21,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q8VIZ0;
Q8VIZ0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
Q8VIZ0
                                                                                                                                                                                                                                    154
                                                                                                                                                                                                                                                                                                                                                 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        896
                                                                                                                112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         354
                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ID DTT DDT DTT DDT DTT DDT DTT DDT DTT DDT DTT DDT DTT DTT
ð
                                                      셤
                                                                                                                à
                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                 ď
                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      음
```

```
37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            63 AGAAAAAAAAAAAAAAAAAAAAGQ-----GGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAAAGGA-G 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GYGGLGSQGTSGP-----GGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGG 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      481 AGGGGYGGL----GSQGTSGPGGYG--PGQQTSG-----RGGLGGQGAGAAAAAAA 526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LGSQGTSGRGCLGGQGAGAAA-----AAAAAAAGGAGQGGYGGLGSQGTSGPG----GY 664
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
               Bishai W.; "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."; Subburated (APR-2001) to the EMBL/GenBank/DDBJ databases. EMBL, AE00163; AAK47971.1; -TIGR; MT3612; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    229 GAGGVGGAGGTGGAGGRAELLFGAGGAGGAGTDGGPGATG--GTGGHGGVGGDGGWLAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  522 GGAGGAGSAGIDATATGATGGTGFSGGAGGAGGATGVGGTNGSGGQGGTGGAGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----GSQGTSGRGGLGGQGAGAAAAAAAAGGAG-QGGYGGLGSQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QGGYGGLGSQ-----GTSGRGGLGG----QGAGAAA-----AAAAAAAGG-AGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AGGVGADNPTGIGGTGGTGGKGGAGGAGAAATATGATGGTGFAGGAGGQGGSSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RGGLGGQCAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              760 VGCINGSG-----GAGGAGGKGGTGGAGGSCADNPIGAGFAGGAGGTGGAAGAGGAGG
                                                                                                                                                                                                                                                                                                  16 GRGGLGGQGAGAAAAAAAAAAAGGAG-QGGYGGLGSQGTSGRGGLGGQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                      211 LGSQGTSGR-----GGLGGQGAGAAAAAAAAAGGAGQGGY----GGLGSQGTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --GISGPGGYGPGQQISGRGGLGGQGA------GAAAAAAAAAAAAAAAAAGGA-GQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GG------LGSQGTSGRGGLGGQ-GAGAAAAAAAAAGGA----GQGGYGGLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     814 TGTGGTGGVVGATGSAGIGGAGGRGDDGASGLGLGLSGFDGGQGGQGGAGGSAGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAGGAGQGGYGGLGSQGTSG---RGGLGGQ-GAGAAA-------AAAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     934 GNGGAGGAGGAGDNNFNGGQGGAGGQGGLGGASTTSINANGGAGGNGGTGGKGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGLGGQCAGAAAAAAAAAAAGG----------AGQGGYGGL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GOGGYGGLGSQG----TSGRGGLGGQGA----GAAAAAAAAAGGA-----GOGGYGG
                                                                                                                                                                                                                                  36.2%; Score 1263.5; DB 16; Length 1715; 40.2%; Pred. No. 1.4e-60;
                                                                                                                                                                                                                                                                   30; Mismatches 271; Indels 206;
                                                                                                            InterPro; IPR000084; PE_region.
InterPro; IPR002173; PfkB.
Pfam; PF0034; PE; 1.
Prodom; PD001223; PE_region; 1.
PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_4.
SEQUENCE 1715 AA; 133832 MW; E85792A774C771B0 CRC64;
                                                                                                                                                                                                                                                 Best_Local Similarity 40.2%
Matches 341; Conservative
                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     329
                                                                                                                                                                                                                                                                                                                                                                                                                                  117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      527
 g
                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
```

σ

```
35.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RV3512 OR MTV023.19.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1089 GGTGGTGGVG 1098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         656 QGTSGPGGYG 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1079 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               053557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DDT T DD DT T DD T 
                                                                                             g
                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                 윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                 ò
                                                                                                                                                   õ
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRAIN-H37RY;

WEDLINE-98295987; PubMed-9634230;

WEDLINE-98295987; PubMed-9634230;

A Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., A Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., A Gordon S.V., Eiglmeier K., Parkhill J., Garnier T., Connor R., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., A Hornsby T., Jagels K., Krogh A., McLean J., Woule S., Murphy L., Oliver S., Osborne J., Quali M.A., Rajandream M.A., Rogers J., A Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; A Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; T. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";

Nature 393:537-544(1998).

REMBL: AL022022: CAA11751.1; -.

RUBERULISE: RV3314; -.

Tuberculist; RV3314; -.

RICHEPTO: IPR000084; PE_region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                s:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SGGAGGOGGAGGAGADNPTGIGGTGGDGGTGGAAGAGGAGGAAGTGGTGGMIGTTGN 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SGRGGLGGQ------GAGAAA-----AAAAAAGGA-------GQGGYGG 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAAAAAAAGGA-GQGGYGGLGSQ-----GTSGRGGLGG----QGAGAAAAAAAG 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   378 AAGNAGVGGAGGQG - GDGGAGGAGADADQPGATGGTGFAGGAGGAGGAGGSSGAGGTNG 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG-LGSQGT 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      249 LGSQG------TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG-LGSQGTSGR 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       556 TGGQGGAGGAGAADNPTGIGGTGGDGGTGGAAGAGGAGGAAGTGGTGGMIGTTGNAGV 615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 260 GHGGVGGDGGWLAPGGAGGAGGQGGAGGAGSDGGALGGTGGTG--GTGGAGGAGGRGALL 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 318 LGAGGOGGLGGAGGOGGTGGAGGDGVLGGVGGTGGKGGVGGVAGLGGAGGAAGQLFSASG 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GAGQGGYGGLGSQGTSGRGGLGGQGA-----GAAAAAAAAAAAAAGGA------GQGG 157
                                                                                                                                                                                                                                           01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
PGRS-family protein.
WY3514 OR MTV023.21.
Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomyceteles; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GRGGLGGQG-----AGAAAAAAAAAAAGGAGQ-----GGYGGLGSQGTSGRGGLGGQGAGA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 16; Length 1489;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             35.9%; Score 1254.5; DB 16; Length 1489
39.5%; Pred. No. 3.7e-60;
ive 36; Mismatches 267; Indels 211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6855CBA1C3CBAF3A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Prodom; PD001223; PE_region; 1.
PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_2.
                                                                                                                                                                                                                         Created)
                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1489 AA; 115754 MW;
                                                                                                                                                                                                                      01-JUN-1998 (TrEMBLrel. 06, 01-JUN-1998 (TrEMBLrel. 06,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 39.59
Matches 336; Conservative
                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00934; PE;
                                              1054 DGGQGGQG 1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Complete proteome. SEQUENCE 1489 AA
672
                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-1773;
665 GPGQQTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                        053559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        296
                                                                                                                         RESULT 13
                                                                                                                                              053559
ð
                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
```

```
481
                                                                                                                                                                                                                                                                                                                                                                                                     505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   564
616 GGAGGQGGDGGAGGAGADADQPGATGGTGFAGGAGGAGKAGGSSSAGGTNSSGSAGGTGR 675
                                                                                                                                                                                                                         735 AGGSSGAGGINGSGGAGGIDGQGGAGGAGADNPTGIGGTGGDGGTGGAAGAGGAGGA 794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-98295887; bubmed-9634230; MEDLINE-98295887; bubmed-9634230; MEDLINE-98295887; bubmed-9634230; MEDLINE-98295887; bubmed-F., Gas S., Barry C.E. III, Tekala F., Gordon S.V., Eighameler K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwall T., Gentles S., Hamlin N., Hollroyd Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares R., Squares R., Stalston J.E., Taylor K., Whitchead S., Barrell B.G.; Deciphering the blology of Mycobacterium tuberculosis from the
                                                                                                                                                                                                                                                                                                                                                                                                                                      SSCAGGINGSGGAGGICGQVVAGGAGISFSNGSNGGIGGIGGVG--GIGGDGGNAGIGAG
                                                                                      506 TSGRGGLGGQGAGAAAAAAAAAAAGGAG-QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               443 AGQGGYGG-LGSQGTSGRGGLGGQ----GAGAAAAAA------AAAAAGGA----
                                                                                                                                                                                                                                                                                                                  000d04-000d0T020CTC2-----C0CCACCTCCCCACCCACCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 565 AAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGA-----GQGGYGGLGSQ
                                                                                                                                                                   386 LGGQGAGAAAAAAAAAAAGGAGGGYGGLGSQGT---SGRGGLGGQGAGAAAAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A79718CDCB74B97D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
PGRS-family protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            complete genome sequence.";
Nature 393:537-544(1998).
EMBL; ALO22022; CAA17749.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81163 MW;
```

DB 16; Length 1079;

Score 1228.5;

```
g
                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
              29;
                                                                                                                                                                   ------GAGAAAAAAAGGAGGGYGGLGSQ----GTSGPGGYGPGQQTSG 180
                                                                                                                                                                                                             181 RGGLGGQGAGAAAAAAAAAAGGAG-QGGYGGLGSQ------GTSGRGGLGGQ 225
                                                                                                                                                                                                                                                                    325
                                                                                                                                                                                                                                                                                                                                            GSQGTSGP---GGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGGGGYGGLGSQGTSG 382
                                                                                                                                                                                                                                                                                                                                                                     510
                                                                                                                                                                                                                                                                                                                                                                                      383 RGGLGGQG------AGAAAAAAAAAGG-AGQGGYGGLGSQ--GTSGRGGLGGQGA 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              542
                                                                                         137
                                                                                                                                    275 TGGTGGNGALGGAGGDAGGSGNSCGNGGLGGKGGNAGAGGAAGSNGGTVGANGTGGD 334
                                                                                                                                                                                                                                                                                                                         451
                                                                                                                                                                                                                                                                                                                                                                                                                                  430 GAAAAAAAAAGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGYGGL 489
                                                                                                                                                                                                                                                                                                                                                                                                                                              571 AGGTGGDRTVGGGTVPAGSGGQG--GNAGGGGAGGQ--GGADGGSGGDGGDAGTGGNGGN 626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGL----GSQGT-SGRGGLGGQGAGAAA 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GTGGDGGLTGTGGTGGSG--GTGGDGGNGGNGADNTANMTAQAGGDGGNGGDGGFGGA 804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  598 AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQG 657
                                                                             66
                                 ASGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 73
                                               ------AGGGGYGGLG------SQGTSGRGGLGGQ
                                                                                                                       GA-GAAAAAAAAAAAGGA-----GQGGYGGLG-----SQGTSGRGGLGGQ-----
                                                                                                                                                                              226 G--AGAAAAAAAAAAAAGAGGGYGGLGSQGTSGRGGLGGQG------AGAAAAAA
                                                                                                                                                                                                                                                                                                  273 AAAAGGAG-----QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGL
                                                                                                                                                                                                                                                                                                                                                           393 VGGAGGAGGSGPNTSPGGNGGQGGGGGGGAG-GAAGAGGANGTAGNGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           G----SQGTSGPGGYGPGQQTSGRGGLGGQGAGAA-AAAAAAAGGAGGGGYG-GLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Arglope aurantia.
Bukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Araneidae; Arglope.
NCBL_TaxID-156844;
Pred. No. 7.2e-59;
5; Mismatches 303; Indels 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       447 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09BIV1;
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequen
01-GCT-2001 (TrEMBLrel. 18, Last annote
Major ampullate spidroin 1 (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
  llarity 40.0%;
Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           658 TSGPGGYGPGQQTSG 672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               805 GAGGGGLTAGANGTG 819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
Best Local Similarity
Matches 318; Conser
                                                                           74 AGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9BIV1
                                  14
                                                                                                                                                                                                                                                                                                                                            326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9BIV1
                                                                                                 8
                                                                                                                                         g
                                                                                                                                                                 ò
                                                                                                                                                                                      合
                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                      ð
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                 ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       240 GAGGGGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              227
                                                                                                                                                                                                                                                                                                                  120 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            420 GRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAA 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34
                                                                                                                                                                                                                                                                                                                                        110 GAGQ-GYG-----SGLGGQGGAGQGGAAAAAAAGGQGGQGGYGGLGSQG----AGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180 GRGGLGGQGAGAAAAAAAAAAAGGGGGCYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          283 QGGQGGYGGLGSQG-AGQGGYGAGQ------GGAAAAAAAAAAAAAAGGAGGEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            540 GSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                             228 GLGGGGGAGGGGAAAAAAAGGGGGHGGYGGLGSQG----AGQGGAGRG-AAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     360 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGAGQGGYGGLGSQGTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     480 GAGOGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     330 G-----XGLGGGGGAGQRGASV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S11k
                                                                                                                                                                                                                                                                              12; Mismatches 103; Indels 197;
                                                                                                                                                                                                                                          34.3%; Score 1199.5; DB 5; Length 447; 50.2%; Pred. No. 1.3e-57;
                                          ;
Spider
                                                                                                                                                                                        447 447
447 AA; 35630 MW; 78D160B7E6DF24C7 CRC64;
                        "Extreme Diversity, runmed-ii.ed.3312;
"Extreme Diversity, Conservation, and Convergence of Fibroin Sequences.";
Science 291:2603-2605(2001).
BMBL, AF350262; AAK30591.1; -.
PRINTS; PR00041; Prion.
PRINTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Search completed: December 18, 2002, 16:20:43 Job time : 48.982 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SAVTSLVSSGGP-----TNPAALSN 436
                   PubMed-11283372;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SQGTSGPGGYGPGQQTSGIRRPAAKLN
                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                           Best Local Similarity
Matches 315; Conserv
SEQUENCE FROM N.A. MEDLINE=21179804;
                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               417
```

7

OM protein

Run on:

Sequence:

Searched:

Database

Result

```
Repetitive protein
SLPIII protein seq
S11k-like protein
Amino acid sequenc
Amino acid sequenc
SLP4 multimeric pr
SLP4 synthetic pr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Spider; orb-weaver; s11k protein; pET((SP1)4/(SP2)1)4; structural protein; purification; fibre; spinning.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Spider recombinant silk protein pET((SP1)4/(SP2)1)4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                         AAY40097
AAU11781
AAU11781
AAY9053
AAY40100
AAY40101
AAB2611
AAB2611
AAB2611
AAB2611
AAB26045
AAY40099
AAR99059
AAR99059
AAR99059
AAR99059
AAR99059
AAR95105
AAR95105
AAR95105
AAR95105
AAR95105
AAR95105
AAR95105
AAR95105
                                                                                                                                                                                                                                                                                                                                                                                                                           AAR41007
AAW53518
AAW53524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAW26348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Butler MM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB32610 standard; Protein; 691 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-2000; 2000WO-US30086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20-JAN-2000; 2000US-0490291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mello CM, Arcidiacono S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MELL/) MELLO C M.
(ARCI/) ARCIDIACONO S.
(BUTL/) BUTLER M M.
(USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2001-483136/52.
N-PSDB; AAH26303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200153333-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26-JUL-2001.
                                         2135
2135
2139
2016
2007
1972:5
1971:5
1961:5
1664:5
1664:5
1668:5
1405:5
13192:5
                                                                                                                                                                                                                                                                                            1346.5
1346.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1144.5
1184.5
1184.5
                                                                                                                                                                                                                                                                                 1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB82610;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
Spider recombinant
Spider recombinant
N. Clavipes spidro
N. Clavipes spidro
N. Clavipes spidro
Dragline protein 1
Dragline protein 1
Dragline protein 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         agiine protein i
clavipes spidro
                                                                                         December 18, 2002, 16:05:10 ; Search time 46.1374 Seconds (without alignments) 1995.696 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                    1. (SIDSZ)gcgdata/geneseqp-embl/AA1980.DAT: *
2. (SIDSZ)gcgdata/geneseqp-embl/AA1981.DAT: *
3. (SIDSZ)gcgdata/geneseqp-embl/AA1981.DAT: *
4. (SIDSZ)gcgdata/geneseqp-embl/AA1981.DAT: *
5. (SIDSZ)gcgdata/geneseqp-embl/AA1981.DAT: *
5. (SIDSZ)gcgdata/geneseqp-embl/AA1981.DAT: *
6. (SIDSZ)gcgdata/geneseqp-embl/AA1981.DAT: *
7. (SIDSZ)gcgdata/geneseqp-embl/AA1981.DAT: *
7. (SIDSZ)gcgdata/geneseqy-embl/AA1981.DAT: *
7. (SIDSZ)gcgdata/geneseqy-embl/AA1981.DAT: *
7. (SIDSZ)gcgdata/geneseqy-embl/AA1981.DAT: *
7. (SIDSZ)gcgdata/geneseqy-embl/AA1991.DAT: *
7. (SIDSZ)gcgdata/gene
                                                                                                                                                US-09-490-291-6
3543
1 MASWIGGQQWGRGSMASGRG......TSGIRIRAPSTSFEHHHHHH 691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
             GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                           Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                   908470 seqs, 133250620 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                   - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB82609
AAM50037
AAM50039
AAM11798
AAU11794
AAU11794
AAU11793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB82610
                                                                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                               A_Geneseq_101002:*
                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              333333355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           691
681
1255
1880
989
1626
1617
818
809
630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0
95.7
68.6
67.3
65.5
64.9
64.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                111.
121.
112.
114.
116.
116.
116.
                                                                                                                                                   Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2431
2431
2384
2321.5
2298
2283.5
2283.5
                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score
```

~

```
ö
                                                                                                                               The present sequence is that of orb-weaver spider (Nephila clavipes)
recombinant silk protein pET((SP1)4/(SP2)1)4. The invention
recombinant silk protein pET((SP1)4/(SP2)1)4. The invention
provides methods for purifying and spinning spider silks and other
structural proteins. Organic acids are used to lyse recombinant;
cells or other biological samples (such as non-recombinant)
cells or other biological samples (such as non-recombinant)
the structural proteins intact. In the case of silk proteins, the
proteins by hydrolysing many of the macromolecules while leaving
the structural proteins intact. In the case of silk proteins, the
cresulting lysate is further purified by ion-exchange or affinity
chromatography and processed into an aqueous-based mixture for
tibre spinning. In the present case, the pET((SP1)4/(SP2)1)4 gene
that condition westor pET34 for recombinant expression in
Escherichia coil, and recombinant silk protein was obtained in
T5-75% purity using propionic acid and anhon-exchange chromatography
to the propionic acid, the recombinant protein was obtained in 80%
purity. Products obtained using the methods of the invention can
to the propionic acid, the recombinant protein was obtained in 80%
purity. Products obtained using the methods of the invention can
to the propionic acid, the recombinant protein was obtained and
parachute cords. The new method has the following advantages
cover prior art: it involves fewer steps, requires less time and
smaller volumes of reagents, results in better recovery of protein
at higher purity (70-99%), is easy to scale up, and the fibres are
spen in an environmentally benign solution reducing hazardous waste
Recovering structural polypeptides in a biological sample, useful for purifying and spinning spider silks and other structural proteins, comprises treating the sample containing the polypeptides with an acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGOGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GYGGLGSQGTSGRGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQT 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAA 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGQGAGAAAAAAAAAAGGGGGGCGCGCGGGGTSGPGGYGPGQQTSGRGGLGGQGAGAAAA 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGAGGGGYGCLGSQCTSGPGCYGPGQCTSGRGCLGGQGAGAAAAAAAGGACGGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361 AAAAAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAAAGGAGQGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 691;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Score 3543; DB 22; 100.0%; Pred. No. 6.3e-246;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                     Claim 2; Page 38-40; 49pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           691; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          accumulation and cost
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             691 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           481
    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
```

540

```
The present sequence is that of orb-weaver spider (Nephila clavipes) recombinant silk protein pOE((SP1)4/(SP2)1)4. The invention provides methods for purifying and spinning spider silks and other structural proteins. Organic acids are used to lyse recombinant calls or other biological samples (such as non-recombinantly proteins by hydrolysing many of the macromolecules while leaving the structural proteins intact. In the case of silk proteins, the resulting lysate is further purified by lon-exchange or affinity chromatography and processed into an aqueous-based mixture for fibre spinning. In the present case, the pOE((SP1)4/(SP2)1)4 gene was cloned into vector pOE-9 for recombinant expression in Escherichia coli, and recombinant silk protein was obtained in 97% purity using propionic acid and anion-exchange chromatography on CAE-Sephadex A50. Products obtained using the methods of the invention can be used in the construction of many materials including films, fibres, woven articles, sutures, ballistic protection, parachutes and parachute cords. The new method has
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Recovering structural polypeptides in a biological sample, useful for purifying and spinning spider silks and other structural proteins, comprises treating the sample containing the polypeptides with an acid
                                                                                                                                                       99
                                                                                                                                   481 GGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGG
                                                       LGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA
                                                                                                               AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Spider; orb-weaver; silk protein; pQE((SP1)4/(SP2)1)4; structural protein; purification; fibre; spinning.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Spider recombinant silk protein pQE((SP1)4/(SP2)1)4.
                                                                                                                                                                                                                 SGPGGYGPGQTSGIRIRAPSTSFEHHHHHH 691
                                                                                                                                                                                         661 SGPGGYGPGQQTSGIRIRAPSTSFEHHHHHH 691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Butler MM;
                                                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 2; Page 34-37; 49pp; English.
                                                                                                                                                                                                                                                                                                                          AAB82609 standard; Protein; 681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-2000; 2000WO-US30086.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20-JAN-2000; 2000US-0490291.
                                                                                                                                                                                                                                                                                                                                                                                                      02-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (MELL/) MELLO C M.
(ARCI/) ARCIDIACONO S.
(BUTL/) BUTLER M M.
(USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mello CM, Arcidiacono
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2001-483136/52.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAH26302.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO20015333-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   26-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                  AAB82609;
                                                                                                               601
                                                                                                                                                                                                                             199
                                                                                                                                                                                                                                                                                     RESULT 2
                                     ò
                                                                         셤
                                                                                                               ò
                                                                                                                                                   a
                                                                                                                                                                                           ò
                                                                                                                                                                                                                               g
```

m

```
Scheller J,
                               13-DEC-2001
                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                   proteins
                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                         55
                                                                                                                                                                                                                                                                                                                                                                                                                                      7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 162
   à
                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                      ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                          ö
                                                                                                                                                                                                                                                                      432
                                                                                                                            192
                                                                                                                                                         430
                                                                                                                                                                                                                                                                                                  492
                                                                                                                                                                                                                                                                                                                 490
                                                                                                                                                                                                                                                                                                                                              550
                                                                                                                     AAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGGGGYGGLGSQGTSG 132
                                                                                                                                                                                 252
                                                                                                                                                                                                             312
                                                                                                                                                                                                                                         372
                                                                                                                                                                                                                                                                                                                                                             612
                                                                                                                                                                                                                                                                                                                                                                                          672
                                                                          Gaps
                                                                                       Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOISOI.
                                                                                                                                                                                                                                                  RGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAG
                                                                                                                                                                              AAAAAAAAAAAGGGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAGGAGQGGYGGLG
                                                                                                                                                                                                             SQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA
                                                                                                                                                                                                                                         AAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                     GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                            the following advantages over prior art: it involves fewer steps, requires less time and smaller volumes of reagents, results in better recovery of protein at higher purity (70-99%), is easy to scale up, and the fibres are spun in an environmentally benign solution reducing hazardous waste accumulation and cost.
                                                                                                                                                                                                                                                                                                                               SQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGGGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                            553 GGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                         613 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQT
                                                                          ö
                                                            Length 681;
                                                                         Indels
                                                          95.7%; Score 3390; DB 22;
100.0%; Pred. No. 5.4e-235;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      clavipes spidroin synthetic homologue SOISO1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAM50037 standard; Protein; 1255 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                  al Similarity 100.
664; Conservative
                                             681 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                |||||
SGIR 674
                                                                                                                                                                                                                                                                                                                                                                                                                      SGIR 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18-SEP-2002
                                                     Query Maccontrol
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        · AAM50037;
                                                                        Matches
                                                                                       13
                                                                                                                    73
                                                                                                                                  71
                                                                                                                                                 133
                                                                                                                                                                              193
                                                                                                                                                                                                            253
                                                                                                                                                                                                                                         313
                                                                                                                                                                                                                                                                                                                                                                                                                      673
                                                                                                                                                                                                                                                                                                                                                                                                                                    671
                                                                                                                                                                                                                                                                     373
                                                                                                                                                                                                                                                                                                                              493
                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 3
 828888
                                                                                        ò
                                                                                                     g
                                                                                                                                 qq
                                                                                                                                                 ò
                                                                                                                                                             g
                                                                                                                                                                                         g
                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AXXXXXXXXXXXXXX
                                                                                                                     à
                                                                                                                                                                           ò
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                      õ
```

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonuclectides, each oligonuclectide encoding a repeating unit of a spidroin protein ripe synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (1) as adhesion surfaces for culturing cells; and (11) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Also they retain water solubility after long-term boiling in aqueous colutions and since they are also soluble in organic solvents but proteinted at high salt concentration, they are easily extracted and purified. The modular construction of the invention facilitates uncorporation of additional peptide-encoding sequences, e.g. to simplify the synthetic N. clavipes spidroin-1 homologue SOISOI described in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  used
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YGGLGSQGTSGRGGLGGQGAGAAA------YGGLGSQGTSGRGAGGGGGGG 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLGSQGT-----SGPGGYGP-GQQTSGRGGLGGQGA 1.91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           243 GLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGA 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gapa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAA------AGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6 GGQQMGRGSMASGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1255;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          68.6%; Score 2431; DB 23; 52.2%; Pred. No. 3.1e-166;
                                                                                                                                                                                                                                                                                               (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----SGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 3.1e-166;
19; Mismatches 39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 22; Page 29-34; 88pp; German.
                                                                                                                                                                                                                                                                                                                                                                Grosse F,
                                                                                                                                 21-MAR-2001; 2001DE-1013781.
                                                                                                                                                                                               39-JUN-2000; 2000DE-1028212.
24-OCT-2000; 2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 52,2%
Matches 585; Conservative
                                                                                                                                                                                                                                                                                                                                                                Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2002-123561/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1255 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; ABL61038
DE10113781-A1
                                                                                                                                                                                            09-JUN-2000;
```

232 AAA
GAGAAAA GAGAAAAAAGGAGGGYGGLGSQGAGRGGQGAGAAAAAAAGGAGGCYGGYGGLGSQGT
GAGAAAAAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGG
AAAAAAGGAGGGYGGLGSQGTSGPGGYGP
GQQTSGRGGLGGQGAGAAAAAAAAGGAGGGGGGGGGGGGGGGGGG
AAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GAGAAAA
LGGQGAGAAAAAAAAAGAGGGGGGGGGGGGGGGGGGGGG
523 AAAAAAAAGGGGGGGGGGGGGGGGG
547
AAAAAAGGGGYGGLGSQGTSGRGGLGGGGAGAAAAAGGAGGAGAAAAAAAAAA
AAAAAAGGAGGYGGLGSQGTSGRGGLGGQGAGAAAA
AAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGG
T 4 039 AAM50039 standard; Protein;
-SEP-2002 (first entry) clavibes spidroin svnthetic homologue SOLSOLSOL
О.
Synthetic.
DE10113781-A1.
13-DEC-2001.
21-MAR-2001; 2001DE-1013781

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidonin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and (iii) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity).

Constitutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and concipitated at high salt concentration, they are easily extracted and concorporation of additional peptide-encoding sequences, e.g. to simplify purification or modulate solubility. This sequence represents the content of construction of the invention described in the content of the invention of 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      34;
                                                                                                                                                                                                                                                   New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               65 AAAGGAGGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----- AAAAAAGGGGGGGGGGGGT---- 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6 GGQQMGRGSMASGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----SGRGGLGGQGAGAAA---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAA------AGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YGGLGSQGTSGRGGLGGQGAGAAA--------AAAAAAGGAGQGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLGSQGT-----SGPGGYGP-GQQTSGRGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       232 AAA-----AAAAAAGGGGYGGLGSQGTSGRGGLGGO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels 478;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1880;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 68.6%; Score 2431; DB 23; Best Local Similarity 52.2%; Pred. No. 4.3e-166; Matches 585; Conservative 19; Mismatches 39;
                                                                                       (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                      Guehrs K;
                                                                                                                                                                                                                                                                                                                                        Claim 22; Page 38-45; 88pp; German.
                                                                                                                                    Grosse F,
                  09-JUN-2000; 2000DE-1028212. 24-OCT-2000; 2000DE-1053478.
                                                                                                                                    Conrad U,
                                                                                                                                                                              WPI; 2002-123561/17.
N-PSDB; ABL61040.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1880 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         266 GAGAAAA----
                                                                                                                                    Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                             proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                162
á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         á
```

```
1010 GGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGSQ 1068
                                                                                                                                                                                                                                                                                 546
                                                                                                                                                                                            463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOISM12.
                                                                        429
                                                                                                                                                             890
                                                                                                                                                                                                                                                                                                                          562
                                                                                                        341 GQQTSGRGCLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGAGAAAA 398
                                                                                                                                                                                                                                      LGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAA 522
                                                                                                                                                                                                                                                                                                                                            891 AGGGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAGGAGGGGGGGGGGGGG 950
GAGAAAAAAGGGGGYGGLGSQGAGRGGQGAGAAAAAAGGAGGGGYGGLGSQGAGRGGO 480
                                                                                                                                                                                                                                                                                           -----SGRGCLGGQGAGAAA-----
                                        481 GAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGQ
                                                                                                                                                  AAAAAAGGAGGGGGGGGGGGGGGGGGT-----SGRGGLGGQ
                                                                                                                                                                                             -----AAAAAAGGAGQGGYGGLGSQGTSGRGG
                                                                                                                                                                                                                                                                              523 AA-------AAAAAAGGAGQGGYGGLGSQGT---------
                                                                                                                                                                                                                                                                                                                                                                                                            -----AAAAAAGGGGGYGGLGSQGTSGRGGLGGQGAGAAA-------

    N. clavipes spidroin synthetic homologue SOISM12 protein.

                                                                                                                                                                                                                                                                                                                      -----SGRGGLGGGGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GAGRGGQGAGAAAAAAGGAGQGGYGGLGSQG-AGRGGQGAG 1108
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------AAAAAGGAGQGGYGGLGSQGTSGPGGYGPG 669
                                                                                                                                                                                                                                                                                                                                                                 ---AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Guehrs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Grosse F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (IPKP-) IPK INST PFLANZENGENETIK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM50038 standard; Protein; 989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000DE-1028212
2000DE-1053478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-MAR-2001; 2001DE-1013781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conrad U,
                                                                                                                                                                                             GAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DE10113781-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAM50038;
                                                                                                                                                                                                                                                                                                                        547
                                                                                                                            601
                                                                                                                                                  399
                                                                                                                                                                      657
                                                                                                                                                                                            430
                                                                                                                                                                                                                 717
                                                                                                                                                                                                                                     464
                                                                                                                                                                                                                                                                                                                                                                                                           601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
                                       g
                                                                                 q
                                                                                                                        g
                                                                                                                                                                      g
                                                                                                                                                                                                             a
                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                             ŏ
                                                                                                      ò
                                                                                                                                                 õ
                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
```

```
This invention describes a novel DNA sequence, encoding synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and artificial organs; (ii) as adhesion surfaces for culturing cells; and cilis as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and clasticity).

Also they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and incorporation of additional peptide-encoding sequences, e.g. to simplify thrifted in organic represents the concentration or modulate solubility. This sequence represents the synthetic N. clavipes spidroin-1 homologue SOISM12 described in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               116 AGRGGQGAGAAAAAGGAGQGGYGGLGSQG-AGRGGLGGGGGAC----AAAAAAAGGAGGGG 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                160 YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAAA-------AA 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6 GGQQMGRGSMASGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAAAGGAGQGGYGGLGSQGT------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---GRGGLGGQGA---GAAAAAAAAAGGAGQGGYGGLGSOGT-----SGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     388 LGGQGAGAAAA-------AAAAAAAGGAGQGGYGGLGSQGTSGRGGLG
                                                                  encoding synthetic spider silk protein, useful e.g. for cl comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        290 QGAGRGGLGGQGAGAAAAAGGAGQGGLGGQGAGAGAAAAAAGGAGQGGYGGLGSQG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SGRGGLGGQGAGAAA-------AAAAAGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 356;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  67.3%; Score 2384; DB 23; 57.8%; Pred. No. 5.8e-163; ive 18; Mismatches 38;
                                                                                                                                    Claim 22; Page 34-38; 88pp; German
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 565; Conservative
                2002-123561/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    989 AA;
                                  N-PSDB; ABL61039
                                                                New DNA encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                  proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   346
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

ø

Thu

```
88888888888888
                                                                                                                                                                          셤
                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                  Dp
                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                            ò
                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                     a
a
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where
                                                                             688
562
                                                                                                                                900
                                                                                                                                                 806
                                                                                                                                                                  622
                                                                                                                                                                                   866
                         487
                                           629
                                                            -----AA 526
                                                                                                               747
                                                                                                                                                                                                                    926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein expression cassette to a green plant.
                                                                                                                                                                                                                                                                                                                                                                     Silk-like protein; SLP; transgenic plant; promoter; 5' terminator; flabitic production; material construction; rope; surgical suture; flabible tie down; electrical component; implantation; bragiline protein 1; DP-1B 16mer; His tag.
                                 GOGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGG
                                                                                                                                                748 AGAAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAAAGGAGQGGYGGLGSQCAGR
                                                                                            AAAAGGAGGGYGGLGSQGTSGRGGLCGQGAGAAA----------
                                                                                                                                                                 -----AAAAAAGGAGGGYGGLGSOGT-----
                                                                                                                                                                                                   -----SGRGGLGGQGAGAAA--------AAAAAAGGAGQGGY
                                                                                                                               --AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAA---
                                                                                                                                                                                                                                                                                                                                                     Dragline protein 1 analogue DP-1B/his tag 16mer.
                                                          YGGLGSQCTSGPGGYGP-GQQTSGRGCLGGQGAGAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  expression cassette to a green plant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 1; Page 87-92; 93pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ε.
                                                                                                                                                                                                                                                                                                 AAU11798 standard; Protein; 1626 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ပ္ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (DUPO ) DU PONT DE NEMOURS &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAY-2000; 2000US-20696BP
                                                                                                                                                                                                                                      699
                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                               GGLGSQGTSGPGGYGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                Nephila clavipes
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200190389-A2
                                                                                                                                                                                                                                                                                                                                    26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                           29-NOV-2001.
                                                                                                                                                                                                                                                                                                                  AAU11798;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            material
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Yang JG;
                                                          488
                                                                            630
                                                                                                              689
                                                                                                                                                                601
                        428
                                         579
                                                                                             527
                                                                                                                               563
                                                                                                                                                                                  807
                                                                                                                                                                                                   623
                                                                                                                                                                                                                    867
                                                                                                                                                                                                                                     653
                                                                                                                                                                                                                                                      927
                                                                                                                                                                                                                                                                                AAU11798
       셤
                        ò
                                        g
                                                          à
                                                                          g
                                                                                             ö
                                                                                                           g
                                                                                                                              õ
                                                                                                                                               g
                                                                                                                                                                ò
                                                                                                                                                                                g
                                                                                                                                                                                                   ð
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                      g
```

```
36;
P is a promoter for driving the expression of an SLP transgene and T is a protein in the method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, blomaterial construction, such as rope, surgical sutures, blomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dragline protein (DP) I synthetic variant DPIB/his tag 16mer (i.e. 16 copies of DPIB with a His tag at the C-terminus) which is used as an SLP construct in the method of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                601
                                                                                                                                                                                                                                                                                                                                                                                  832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         505
                                                                                                                                                                                                                                                                                                        888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            943
                                                                                                                                                                                                                                                                                                                                                    97
                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                      11 GRGSMASGRGGLGGQGAGAAAA-------AAAAAAGGAGQGGYGGL 49
                                                                                                                                                                                                                                                                                                                                                  889 GGGGAG----AAAAAAGGAGGGYGGLGSQG-AGQGGYGGLGSQGAGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              361 -AAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGGGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          T-------AAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQTSGRGGLGGQGAGAAAA------AAAAAAGGAGGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAAAAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAA-----
                                                                                                                                                                                                                           DB 23; Length 1626;
                                                                                                                                                                                                                        65.5%; Score 2321.5; DB 23; Lengt 57.6%; Pred. No. 2.5e-158; ive 14; Mismatches 51; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QGAGAAAA-----
                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                            1626 AA;
                                                                                                                                                                                                                                        Local Similarity
es 556; Conserv
                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                          Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1056
```

7

```
322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     681
                                                                                                                                                                                                                                                                                                                                                                                      361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAU1179.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                         à
                                              원
                                                                          ò
                                                                                                   g
                                                                                                                            ò
                                                                                                                                                  셤
                                                                                                                                                                               ö
                                                                                                                                                                                                   셤
                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassetter. The green plant contains the expression cassetter P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, flexible its downs for certain electrical components, or as a biomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dagline protein (DP) I synthetic variant DPLB Ismer (i.e. 16 copies of DPLB) which is used as an SLP construct in the method of the invention.
                                                                                                                                                                                                                                                                                            Orb-weaving spider; silk-like protein; SLP; transgenic plant; prometer; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation; Dragline protein 1; DP-1B 16mer.
             Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein
                                                 1562 GSQGAGQGAGAAAAAGGAGGGGGGGGGGGGG-AGRGGQGAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64.9%; Score 2298; DB 23; Length 1617; 61.5%; Pred. No. 1.2e-156; 1ve 13; Mismatches 39; Indels 288;
AAAAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAA-
                                                                                                                                                                                                                                                                  Dragline protein 1 analogue DP-1B 16mer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        expression cassette to a green plant -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Page 77-81; 93pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ε.
                                                                                                                                                                                           Æ
                                                                                                                                                                                        AAU11794 standard; protein; 1617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (DUPO ) DU PONT DE NEMOURS & CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAY-2000; 2000US-206968P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2002-106209/14.
                                                                                                                                                                                                                                           (first
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1617 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                |||||||
НИНННН 1626
                                                                                                                                                                                                                                                                                                                                                          Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                               WO200190389-A2.
                                                                                                                                                                                                                                           26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Simi
hes 544;
                                                                                                                                                                                                                                                                                                                                                                         Synthetic.
                                                                                                                                                                                                                   AAU11794;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Yang JG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loca
Matches
                                                                                                  989
                                                                                                                          1621
                                                 639
                                                                                                                                                                  RESULT 7
                                                                                                                                                                               AAU1179
                                                                       셤
                   g
                                                 ò
                                                                                                   ö
                                                                                                                        셤
                                                                                                                                                                                           ò
```

36;

--- AAAAAAGGAGQGGYGGL 49

GRGSMASGRGGLGGQGAGAAA--

11

ŏ

Conservative

```
72 GSQG----AGRGGQGAG----AAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGRGG 123
                                                                                                        124 LGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGR 179
                                                                                                                                                                                                       510 AGGGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAAGGAG 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              570 QGGYGGLGSQG----AGRGGQGAG----AAAAAAGGAGQGGYGGLGSQG-AGQGGYGGLG 620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation;
GLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGL
                                                                                                                                                                                                                                                           196 A--------AAAAAGGAGGKGGLGSQGTSGRGCLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                          QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       593 QGAGAAAA--------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQ
                                  GSQCTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQCT-----SGRGG
                                                                                                                                                                                   137 GGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                   235 AAAAAAGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                            273 ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    506 QQTSGRGGLGGQGAGAAAA------AAAAAAGGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              545 GTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dragline protein 1 analogue DP-1B/his tag 8mer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAU11797 standard; Protein; 818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAU11797;
```

360

436

610 493

531

584 721

781

```
Orb-weaving spider; silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation; bragline protein 1; DP-1B 8mer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SQCTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGIRIR 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein expression cassette to a green plant -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QGGLGSQGAGGAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AAAAAAGG
                                                                                                                                                                                                 AGOGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQ
                                                                                                                                                                                                                                                       QGTSGPGGYGP-GQQTSGRGCLGGQGAGAAAA------AAAAAAG
                                                                                                                                                                                                                                                                                                                             GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                 GAGOGGYGGLGSQG----AGRGGQGAG----AAAAAAGGAGQGGYGGLGSQGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----SGRGGLGGQGAGAAA--------AAAAAAGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGS
                      AAAAAAAAGGAGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGCQGAGAAA-
                                                                                                                                                               AGQGGYGGLGSQGT-----SGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Dragline protein 1 analogue DP-1B 8mer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ξ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAU11793 standard; protein; 809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (DUPO ) DU PONT DE NEMOURS & CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-MAY-2000; 2000US-20696BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -------ннннн 818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                679 АРЅТЅFЕННИНН 691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200190389-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               29-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAU11793;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yang JG;
                                                                                                                                                                                                                                                                                                                                            611
                    309
                                                     390
                                                                                                                          447
                                                                                                                                                                                                 499
                                                                                                                                                                                                                                    437
                                                                                                                                                                                                                                                                                                                                                                              532
                                                                                                                                                                                                                                                                                                                                                                                                                 670
                                                                                                                                                                                                                                                                                                                                                                                                                                                    585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               813
                                                                                                                                                             407
                                                                                                                                                                                                                                                                                                         484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAU11793
                                                                                                                      g
                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                             ò
                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, flexible tie downs for certain electrical components, or as a biomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dragline protein (DP) I synthetic variant DPIB/his tag Merc (i.e. 8 copies of DPIB with a His tag at the C-terminus) which is used as an SLP construct in the method of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 AAGGAGGGYGGLGSQG----AGRGGQGAG----AAAAAGGAGGGGYGGLGSQGAGGG 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----SGRGCLGGQGAGAAAA-------AAAAAAGGAGQGGYG 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAA 308
                                                                                                                                                                                                                                                                                                                                     Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein expression cassette to a green plant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MASMTGGQQMGRGSMAS---GRGGLGGQGAGAAA----------AAAA 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----AAAAAGGAGGGYGGLGSQGTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---SGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                37 AAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 23; Length 818;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41; Indels 317;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64.5%; Score 2283.5; DB 2: 59.6%; Pred. No. 7.5e-156;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GCLGGQGAGAAAAAAAAGGAGQCGYGGLGSQGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11; Mismatches
8mer; His tag.
                                                                                                                                                                                                                                                                                                                                                                            expression cassette to a green plant
                                                                                                                                                                                                                                                                                                                                                                                                             Example 1; Page 83-86; 93pp; English.
                                                                                                                                                                                                                                  ΕΙ.
                                                                                                                                                                                                                               (DUPO ) DU PONT DE NEMOURS & CO
                                                                                                                                                            24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                             25-MAY-2000; 2000US-206968P
Dragline protein 1; DP-1B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                     WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     818 AA;
                                 Nephila clavipes.
                                                                                     WO200190389-A2.
                                                                                                                        29-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         544;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Seguence
                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                                             material
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                    Yang JG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               336
```

à g à g ò 셤 õ 윱 ð g g ð g

à

G----AGRGGQGAG----AAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGG 732

681

Thu Dec 19 09:28:21 2002

```
g
                                                            õ
                                                                              ద
                    ò
                                                                                                                                                                                                                 34;
                                    in a green plant by providing a green plant containing an SIP expression cassette. The green plant contains the expression cassette P-SIP-T where p is a prometer for driving the expression of an SIP transgene and T is a 5' terminator. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in flabiles or in material construction, such as rope, surgical sutures, blomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dragline protein (DP) I synthetic variant DPIB Bmer (i.e. 8 copies of DPIB) which is used as an SIP construct in the method of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       398
                                                                                                                                                                                                                                                                                                                               286
                                                                                                                                                                                                                                                                                                                                                                                                                                                       360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRGGLGG 592
                                                                                                                                                                                                                                                                                                                   LGGQGAGAAAA--------AAAAAAGGAGQGGYGGLGSQGTSGRGGL 136
                                                                                                                                                                                                                                                                                                                                                            137 GGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGLGGQGAGAAA 195
                                                                                                                                                                                                                                                                                                                                                                      T------AAAAAGGAGAGAAAA-------AAAAAAGGAG 446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   447 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-G 505
                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                              72 GSQG----AGRGGQGAG----AAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGRGG 123
                                                                                                                                                                                                                                                                                                                                                                                                     -----AAAAAGGGGGGGGGGGGGGGGGGGGGGGGGAAAA 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----AAAAAAGGAGQGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           506 QQTSGRGGLGGQGAGAAA-------AAAAAAGGAGGGYGGLGSQ 544
                                                                                                                                                                                                                                    GRGSMASGRGGLGGQGAGAAA-------AAAAAAGGAGQGGYGGL 49
                                                                                                                                                                                                                                               -----SGRGG 97
                                                                                                                                                                                                                                                                                                                                                                                                               -AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                           235 AAAAAAGGAGQGGYGGLGSQGT-----SGRGGLGGOGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-------
                                                                                                                                                                                                                 Indels 286;
                                                                                                                                                                                           63.6%; Score 2254; DB 23; Length 809; 61.6%; Pred. No. 9.5e-154; ive 11; Mismatches 35; Indels 286,
                             The invention relates to a method of producing silk-like
                                                                                                                                                                                                                                                                           GSQCTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQCT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Page 74-76; 93pp; English.
                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                    196 A-----
                                                                                                                                                                        809 AA;
                                                                                                                                                                                                              533;
                                                                                                                                                                          Seguence
                                                                                                                                                                                            Query Match
                                                                                                                                                                                                        Local
                                                                                                                                                                                                      Best Loca
Matches
                                                                                                                                                                                                                                                                            20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   545
g
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
squentially arranged oligonucleotides, each comprising a group of a sequentially arranged oligonucleotides, each comprising a group of a sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover attifictal organs; (11) as adhesion surfaces for culturing cells; and cilii as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Chiso they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high sait concentration, they are easily extracted and purified. The modular construction of the invention facilitates in concentration or additional peptide-encoding sequences, e.g. to simplify purification or modulate solubility. This sequence represents the synthetic N. clavibes spidroin-1 homologue Sol described in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This invention describes a novel DNA sequence, encoding a synthetic
                                  -----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 178;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          filter;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 630;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Spidroin; spider; silk; fibre; film; membrane; wound;

    N. clavipes spidroin synthetic homologue SO1 protein.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60.7%; Score 2151; DB 23;
68.6%; Pred. No. 1.8e-146;
11ve 9; Mismatches 41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Guehrs K;
                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 22; Page 48-50; 88pp; German.
                                                                                                                         929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Grosse F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (IPKP-) IPK INST PFLANZENGENETIK
                                                                                                                                                                                                                                                                                                                             AAM50042 standard; Protein; 630
                                                                                                                09-JUN-2000; 2000DE-1028212.
24-OCT-2000; 2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-MAR-2001; 2001DE-1013781
                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
QGAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2002-123561/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    630 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; ABL61043.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DE10113781-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    .8-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 497;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               invention
                                                                                                                                                                                                                                                                                                                                                                                            AAM50042;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local
                                                                                                                      632
593
                                                                                                                                                                              789
                                                                                                                                                                                                                                                                        RESULT 10
                                                                                                                                                                                                                                                                                                        AAM50042
```

```
360
                                             125
                                                     126 GSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGG 185
                                                                                                 186 LGGQGAGAAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 236
                                                                                                                                                      151 LGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQG-AGRGGLGGQGAG---AA 205
                                                                                                                                                                           237 AAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSG 296
                                                                                                                                                                                                    248
                                                                                                                                                                                                                     297 RGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGA 355
                                                                                                                                                                                                                                  356 GAAAA--------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGA 395
                                                                                                                                                                                                                                                                                                             409
                                                                                                                                                                                                                                                                                                                                                          515
                                                                                                                                                                                                                                                                                                                                                                   909
                                                                                                                                                                                                                                                                                                                                                                                                                                                          567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                611
  65
                       9
                                                                                                                                                                                      GGOOMGRGSMASGRGGLGGQGAGAAAAAAAAGGAGGGGCYGGLGSOGTSGRGGLGGGGA
                                                                                                                                                                                                                                                                                                                                                         QGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                             568 -AGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAA--------AAAAAA
                    9 GGQGAGQG----GYGGLGGQGAGQGAGAAAAAGGAGQGGYGGLGSQG----AGRGGQGA
                                        66 GAAAAAAAAAAAAAAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGAGQGGYGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT--SGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361 --- AAAAAAGGAGGGYGGLGSQG----AGRGGQGAG----AAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                    516 GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Spidroin; spider; silk; fibre; film; membrane; wound; filter; SO1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Synthetic spidroin homologue SO1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N. clavipes spidroin synthetic homologue SOI protein #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1..28
/label= LeB4_signal_peptide
29..659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM50047 standard; Protein; 676 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note= "S
660..672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GYGPG 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          612 GQGAG 616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAM50047;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    999
                                                                                                                                                                                                                                                                                                                                                        456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAM50047
                                                              셤
                                                                                                         a
                                                                                                                                                   셤
                                                                                                                                                                                             g
                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                       පු
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              අ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                          à
                                                                                    ò
                                                                                                                             ò
                                                                                                                                                                           ò
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                  ç
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
spider all protein, comprising and addition as a sequencially a synthetic speak are accomplished, each comprising a synthetic sequentially arranged oilgonucleotides, each oilgonucleotide encoding a repeating unit of a spidonucleotide, each oilgonucleotide encoding a repeating unit of a spidonucleotide, each oilgonucleotide encoding a repeating unit of a spidonucleotide. The synthetic protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) con medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and artificial organs; (ii) as adhesion surfaces for culturing cells; and stifficial organs; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity).

Also they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and precipitated at high salt concentration, they are easily extracted and composed of the Lebd signal peptide encoding sequences, e.g. to simplify purification or modulate solubility. This sequences represents a construct composed of the Lebd signal peptide, N. clavipes spidroin-I synthetic homologue SOl, a c-Myc-tag and an endoplasmic reticulum (ER)-retention signal described in the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         32;
                                                                                                                                                                                                                                                                                                                                                                          DNA encoding synthetic spider silk protein, useful e.g. for closing ids, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     126 GSQCTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGG 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This invention describes a novel DNA sequence, encoding a synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6 GGQQMGRGSMASGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGYGGLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  179 LGGOGAGAAAAAGGAGOGGYGGLGGQGAGOGGYGGLGSQG-AGRGGLGGQGAG----AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 178;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 676;
                                                                                                                                                                                                                                                        (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              60.7%; Score 2151; DB 23;
68.6%; Pred. No. 1.9e-146;
/note= "c-myc-tag"
673..676
/note= "ER retention signal"
                                                                                                                                                                                                                                                                                               Grosse F, Guehrs K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 1; Fig 10A; 88pp; German.
                                                                                                                                                        21-MAR-2001; 2001DE-1013781.
                                                                                                                                                                                             2000DE-1028212.
2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 497; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                186 LGGQGAGAAAAAAAA---
                                                                                                                                                                                                                                                                                             Conrad U,
                                                                                                                                                                                                                                                                                                                                   WPI; 2002-123561/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         676 AA;
                                                                             DE10113781-A1
                                                                                                                                                                                               09-JUN-2000;
                                                                                                                                                                                                                   24-OCT-2000;
                                                                                                                                                                                                                                                                                               Scheller J,
                                                                                                                  13-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                               proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                      Domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         237
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
```

RGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGA 355

297

à

```
Nephila clavipes; silk protein; tandem repeat; fibre; dragline; tensile strength; elasticity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The present sequence represents a spider silk protein from the present invention. Spider silk proteins, and peptide fragments within the proteins, can be produced and purified independently and can then be mixed and made into fibres that have higher tensile strengths and
                                                                                          515
395
                                                          455
                                                                                                   528
                                                                                                                                                           909
                                                                 567
                                                                                                                                                                          587
                                                                                                                                                                                           664
                                                                                                                                                                                                           639
                       568 -AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA--------AAAAAA
                                                                                                                                                                  QCTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLG
                                                                                                                          GOGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               fibres
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Recombinant spider silk proteins - useful for making
                                                                                                                                                                                                                                                                                                                                    Nephila clavipes spider silk protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 1; Column 29-34; 68pp; English.
                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                   AAW53346 standard; Protein; 718
                                                                                                                                                                                                                                                                                                                                                                                                                                            91US-0684819.
90US-0511792.
94US-0317844.
95US-0425069.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Xu M;
                                                                                                                                                                                                                                                                                                                                                                                                                             95US-0425069
                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hinman MB, Lewis RV,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (UYWY-) UNIV WYOMING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1998-270437/24.
                                                                                                                                                                                                                                                                                                                                                                           Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAV23249
                                                                                                                                                                                                                           GYGPG 669
                                                                                                                                                                                                                                          GOGAG 644
                                                                                                                                                                                                                                                                                                                                                                                                                             19-APR-1995;
                                                                                                                                                                                                                                                                                                                    06-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-APR-1990;
04-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                            US5728810-A
                                                                                                                                                                                                                                                                                                                                                                                                                                            15-APR-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-APR-1995
                                                                                                                                                                                                                                                                                                                                                                                                            17-MAR-1998
                                                                                                                                                                                                                                                                                                    AAW53346;
                                                                                                                                                                                                                                                                                                                                                    Spider;
                                                                                                                                                                                                                                                                                                                                                             coccon;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB;
        277
                         356
                                         331
                                                                                         456
                                                                                                        438
                                                                                                                         516
                                                                                                                                         477
                                                                                                                                                                          529
                                                                                                                                                                                                         588
                                                                                                                                                                                                                          999
                                                                                                                                                                                                                                          640
                                                                                                                                                                                          607
                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                            AAW53346
                        ò
                                       쉽
                                                        à
                                                                       ď
                                                                                     ŏ
                                                                                                       g
                                                                                                                        ò
                                                                                                                                       셤
                                                                                                                                                        ò
                                                                                                                                                                        g
                                                                                                                                                                                          ò
                                                                                                                                                                                                        g
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                         a
```

```
35;
elasticity than naturally occurring fibres. The fibres can be used in mixed composites. The invention allows the two naturally occurring Nephila clavipes silk proteins to be produced independently so that they can later be combined to form silk fibres of high tensile strength
                                                                                                                                                                             107
                                                                                                                                                                                                                                                                             291
                                                                                                                                                                                                                                                                                                260
                                                                                                                                                                                                                                                                                                                    350
                                                                                                                                                                                                                                                                                                                                      311
                                                                                                                                                                                                                                                                                                                                                                                                452
                                                                                                                                                                                                                                                                                                                                                                                                                   419
                                                                                                                                                                                                                                                                                                                                                                                                                                      499
                                                                                                                                                                                                                                                                                                                                                                                                                                                          475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            636
                                                                                                                                                                                               201
                                                                                                 Gaps
                                                                                                                                      26
                                                                                                                    81
                                                                                                                                                                                                                                    25 QGAGAAAAAAAAAAGGAGGGGGGGGGGGGGGT--SGRGGLGGGGGAG-AAAAAAAAAAAAAAAA
                                                                                                                                     1 QGAG----AAAAAAGGAGGGYGGIGGQGAGQGGYGGIGGQGAGQGAGAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                  QGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                      LGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT-----SGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GCQCAGAAAA-------AAAAAAGCAGQCCYGCLGSQGTSGRGCLGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              593 QGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGCQCAGAAAAAAAGGACQCGY
                                                                                                                                                                                                                                                                            SQGTSGRGCLGGQGAGAAAAAA---------AAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                         351 GGQGAGAAAA------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGGG
                                                                                                                                                                                                                                                                                                                                                                                               AGAAAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGGGAGAAAAAAAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           588 QGAG----AAAAAAGGAGGGYGGLGGQGV-GRGGLGGQGAG-----AAAAGGAGQGGY
                                                                                                 Indels 184;
                                                                              Length
                                                                           60.3%; Score 2138; DB 19;
66.3%; Pred. No. 1.7e-145;
iive 20; Mismatches 49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGLGSQGTSGPGGYGPGQQTSGIRIRAPSTS 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GUGSGASAA-----SAAASRLSSPQAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                 tches 498; Conservative
                                                                                       Similarity
                                                          718 AA;
                                        and elasticity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-MAR-2000
                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY59070
                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY59070;
                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                                                                     453
                                                                                                                                                         8
                                                                                                                                                                                                                                                        143
                                                                                                                                                                           57
                                                                                                                                                                                                                  108
                                                                                                                                                                                                                                                                                                                                      261
                                                                                                                                                                                                                                                                                                                                                                                               393
                                                                                                                                                                                                                                                                                                                                                                                                                  371
                                                                                                                                                                                                                                                                                                                                                                                                                                                        120
                                                                                                                                                                                              142
                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                                                                                               202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           200
                                                                                                                                                                                                                                                                           253
                                                                                                                                                                                                                                                                                                                  292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY59070
ID AAY5
XX
AC AAY5
XX
DT 08-N
 888888
                                                                                                                     δ
                                                                                                                                    q
                                                                                                                                                        ò
                                                                                                                                                                          a
                                                                                                                                                                                              ò
                                                                                                                                                                                                                g
                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

Page 12

```
Cosmetic or dermatological composition containing spider silk protein, for hair or skin care, in make-up or sunscreens
                      The present sequence represents the natural spider silk protein spidroine major 1. The protein improves the moisturizing/softening action of the compositions. The protein, and its fragments are used in cosmetic or dermatological compositions. These compositions have use as hair or skin care products; and make-up or sunscreens. As the protein is a good, persistent film-formers on the skin of low surface density, it can be used for delivery of active agents that are generally difficult to administer, e.g. vitamins, hormones, moisturizers or agents for treating disorders of the
                                                                                                                                                                                                                          GGYGP-GQQTSGRGGLGGQGAGAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                        ----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                     532 GGQGAGAAAAAAGGAGQGGLGGQGAGQGAGAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                        QGAGAAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 176;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Spider silk protein; spidroine major 1; cosmetic; make-up; dermatological compositions; hair care; skin care; sunscreen; hormone; moisturizer; skin disorder; skin disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 651;
     LGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     60.3%; Score 2135; DB 20;
68.3%; Pred. No. 2.5e-145;
tive 15; Mismatches 39;
                                                                                                                                                                                                                                                                         653 GGLGSQGTSGPGGYGPGQQTSGIRIRAPSTS 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Spider silk protein spidroine major 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Arrandeau JP;
                                                                                                                                                                                                                                                                                              637 GGVGSGASAA-----SAAASRLSSPQAS
                                                                                                                                                                                                                                                                                                                                                                                               Ą.
                                                                                                                                                                                                                                                                                                                                                                                          AAY40097 standard; protein; 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 3; Fig 1; 32pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98FR-0001614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98FR-0001614
                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 68.38
Matches 495; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Philippe M, Garson JC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1999-510729/43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (OREA ) L'OREAL SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        651 AA;
                                                                                                                                        GGGGAGAAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                             AAY40097;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      skin and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
   453
                                  420
                                                                    200
                                                                                                   94
                                                                                                                                    553
                                                                                                                                                                                                      593
                                                                                                                                                                                                                                                                                                                                                                              AAY4009
                                 g
                                                                  ò
                                                                                                   8
                                                                                                                                    ò
                                                                                                                                                                     g
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                           35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       proteins. The spider silk proteins are characterized by repeting alpha and beta regions and optional variable regions. The DNA sequences are useful in the production of spider silk protein by recombinant DNA techniques. The recombinant place silk proteins may be used for the production of fibers. The present sequence represents the spider silk protein 1, derived from the major ampullate gland of Nephila clavipes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             spider silk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              202 AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA.----AGGAGQGGYGGLG 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---AAAAGGAGQGGYGGLGS 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              202 SQG-AGRGGLGGQGAGAAAAAAAGGAGQGGLGGQCAGQGAGASAAAAGGAGQGGYGGLGS 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82 GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 141
                                                                                                                                                                                                                                                                                                                                                                       Isolated DNA, vector and transformed cell encoding for and useful in the production of spider silk protein \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Spider silk protein; dragline silk protein; major ampullate gland;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60.3%; Score 2138; DB 21; Length 718;
66.3%; Pred. No. 1.7e-145;
1.ve 20; Mismatches 49; Indels 184;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       The invention provides isolated cDNA molecules coding for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOCTSGRGGLGGQGAGAAAAAA
N. clavipes spider silk protein 1.
                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 1; Fig 6A-D; 65pp; English.
                                                                                                                                                                                                                                                                                       Hinman MB, Xu M, Lewis RV;
                                                                                                                                                                                                   91US-0684819.
                                                                                                                                                                  94US-0317844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 498; Conservative
                                                                                                                                                                                                                                                      (UYWY-) UNIV WYOMING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                                                                                                                                                                                                                                                                                    WPI; 2000-061225/05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          718 AA;
                                                                Nephila clavipes
                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAZ38195
                                                                                                                                                                  04-OCT-1994;
                                                                                                                                                                                                   .5-APR-1991;
                                                                                                                                                                                                                    20-APR-1990;
                                                                                               JS5989894-A
                                                                                                                                  23-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   371
```

ð 셤 ð 셤 ò a 셤

õ

유

à

a

å 윰 å

å

34;

Gaps

24-MAY-2001; 2001WO-US16937

```
Orb-weaving spider; spidroin 1; silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation.
                                                                                                                                                                                                        392
                                                                                                                                                                                                                                                                                                             552
                                                                                                                                                                                                                                                                                                                                                                                                989
                               350
                                                                                                                                                                                                                  592
                                                                                                          291
                                                                                                                                                                                       311
                                                                                                                                                                                                                                          393 AGAAAAAAAAAAGGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG 452
                                                                                                                                                                                                                                                                            LGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGAGGGGYGGLGSQGT-----SGP 499
                                                                                                                                                                                                                                                                                                                             531
                                                                                                                                                                                                                                                                                                                                                              587
                                                                                                                                                                                                                                                                                                                                                                               652
                                                                  G---AAAAAAAGGAGGGYGGLGNQG------AGRGGQG------AAAAA 142
        351 GGQGAGAAAA------AAAAAAGGAGQGGYGGLGSQGTSGRGCLGGQG
                                                                                                                                                                                                                                                                                                                     QGAGAAAAAAAAAGGAGGGGYGGLGSQGT--SGRGGLGGQGAG-AAAAAAAAAAAGGAGQ
                                                                                                   SQGTSGRGGLGGQGAGAAAAAA-------AAAAGGAGQGGYGGLGS
                                                                                                                                                                      QGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                                             GGYGP-GQQTSGRGGLGGQGAQAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                              GGQGAGAAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                                                                                                                                                                                               QGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Spider natural silk protein Spidroin 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAU11781 standard; protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                 GGLGS 657
                                                                                                                                                                                                                                                                                                                                                                                                                                 GGVGS 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAU11781;
                                                                  142
                                                                                 108
                                                                                                                                    253
                                                                                                                                                     202
                                                                                                                                                                      292
                                                                                                                                                                                       261
                                                                                                                                                                                                                         312
                                                                                                                                                                                                                                                          371
                                                                                                                                                                                                                                                                           453
                                                                                                                                                                                                                                                                                           420
                                                                                                                                                                                                                                                                                                            500
                                                                                                                                                                                                                                                                                                                            476
                                                                                                                                                                                                                                                                                                                                             553
                                                                                                                                                                                                                                                                                                                                                              532
                                                                                                                                                                                                                                                                                                                                                                               593
                                                                                                                                                                                                                                                                                                                                                                                               588
                                                                                                                                                                                                                                                                                                                                                                                                                 653
                                                                                                                                                                                                                                                                                                                                                                                                                                 637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAU11781
ID AAU
                                               g
                                                                             g
                                                                                                                  g
                                                                                                                                                 g
                                                                                                                                                                                   g
                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                  ò
                                                                                                  õ
                                                                                                                                ð
                                                                                                                                                                     ò
                                                                                                                                                                                                    οχ
                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                           ď
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing sliks and slik-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, flexible tie downs for certain electrical components, or as a biomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is orb-weaving spider spidroinl, a natural silk protein whose sequence is used as the basis for the SLP proteins used in the method of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  G---AAAAAAAGGAGGGGGGGGGGGG----AAAAAA 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AGAAAAAAAAAGGAGGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAAAGGAGQGGYGG 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LGSQGTSGRGGLGGQGAGAAAAAAAAGGAGGGYGGLGSQGT-----SGP 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGYGP-GQQTSGRGGLGGQGAQAAAAAAAA----AGGAGQGGYGGLGSQGTSGRGGL 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25 QGAGAAAAAAAAAGGAGGGGGGGGGGGGT--SGRGGLGGQGAG-AAAAAAAAAAAAGGAGQ 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAA------AGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              143 AGGAGOGGYGGLGSQG-AGRGGLGGQGAGAAAAAGGAGQGGYGGLGGQGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SQGTSGRGGLGGQGAGAAAAAA-------AAAAGGAGGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                292 QGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       expression cassette to a green plant, useful in fabrics expression cassette to a green plant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels 176;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 651;
                                                                                                                                                                                                                                                                                                                                                                                                                                             60.3%; Score 2135; DB 23;
68.3%; Pred. No. 2.5e-145;
ive 15; Mismatches 39;
                                                                                                                                                                                          Disclosure; Page 69-71; 93pp; English.
                                                      н
                                                      ы
                                                      ႘
                                                   (DUPO ) DU PONT DE NEMOURS &
                         25-MAY-2000; 2000US-206968P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 495; Conservative
                                                                                                          WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                     651 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                  Yang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         200
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
```

Search completed: December 18, 2002, 16:17:19 Job time: 53.1374 secs

Sequence (Sequence (Sequen

sednence sed

Sequence

```
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Xu, Ming
APPLICANT: Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CEL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
CORRESPONDENCE: 69
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 184; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 718;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONTEXPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5728810th Washington Street
CITY: Falls Church
STATE: 101 No. 5728810th Washington Street
STATE: 20046
CONFIRY: U.S.A.
ZIP: 22046
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATCHIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425,069
FILING DATE: 19-APR.1995
CLASSIFICATION: 435
ATTONNEY/AGENT INPORMATION:
NAME: MULPHY JT., GGTAIG M
REGISTRATION NUMBER: 28,977
REBERENCE/POCKET NUMBER: 28,977
RELECOMMULICATION INFORMATION:
TELLEPHONE: (703) 205-8050
THELEFAX: (703) 205-8050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60.3%; Score 2138; DB 1;
66.3%; Pred. No. 3.8e-147;
iive 20; Mismatches 49;
US-08-477-509B-83
US-08-482-085B-83
US-08-844-791A-83
US-08-425-069-4
US-08-317-844B-4
US-08-475-411A-36
US-08-475-411A-36
US-08-475-411A-41
US-08-475-411A-41
US-08-475-411A-41
US-08-475-618-89
US-08-4482-085B-89
US-08-444-791A-89
                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-175-155-54
US-08-089-862-11
US-08-587-333-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 2, Application US/08425069
Patent No. 5728810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 66.3%;
Matches 498; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           718 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
     11101
11101
1101
1101
138
1038
1038
11332
1332
1332
1332
1018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
US-08-425-069-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; MOLECULE TY
US-08-425-069-2
  1184.5
1184.5
1184.5
1187.5
1138.5
1138.5
1138.5
1137.5
1137.5
1137.5
1135
1135
1135
1135
1135
1135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
     8001284888888444444
800128488888844444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appli
                                                                                                                                                                                                                            (without alignments)
1277.936 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                              December 18, 2002, 16:14:14; Search time 15.9094 Seconds
                                                                                                                                                                                                                                                                                                                                                             MASMIGGOOMGRGSMASGRG.....TSGIRIRAPSISFEHHHHHH 691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 2 Sequence 2 Sequence 2 Sequence 2 Sequence 2 Sequence 6 Sequence 6 Sequence 6 Sequence 6 Sequence 7 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                             GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-317-8448-2

US-08-317-8448-2

US-09-034-177-3

US-09-247-806-1

US-09-247-806-4

US-09-247-806-4

US-09-247-806-6

US-09-247-806-8

US-09-247-806-8

US-09-247-806-8

US-09-247-806-9

US-08-478-806-10

US-08-478-806-10

US-08-478-806-10

US-08-478-806-10

US-08-478-806-10

US-08-478-806-10

US-08-478-806-10

US-08-478-806-10

US-08-478-806-10

US-08-478-808-84

US-08-478-808-84

US-08-478-918-64

US-08-478-918-64

US-08-478-918-64

US-08-478-918-64

US-08-478-918-64

US-08-478-918-64

US-08-478-918-64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Potal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         262574 segs, 29422922 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                         - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                        US-09-490-291-6
3543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2138
2135
2135
2135
22135
2076
2076
1972.5
1971.5
1405
1346.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
1241.5
124
                                                                                                                                                                                                                                                                                                                                  Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                       Scoring table:
                                                                                                                                         OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                                                                                                                                                 Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Result
```

35;

81

```
g
                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                  엄
                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Lewis, Randolph V.
APPLICANT: Lewis, Ming
APPLICANT: Xu, Ming
APPLICANT: Xu, Ming
APPLICANT: Stupent Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSE: Birch, Stewart, Kolasch & Birch
                                                                                                                                                                           392
                                                                                                                                                                                                                                                          452
                                                                                                                                                                                                                                                                  108 G---AAAAAAAGGAGGGYGGLGNQG------AGRGGQG------AAAAA 142
                                                                                                                                                                                                                                                                                               531
                                                                                                                                                                                                                                                                                                                                                                                                                201
                                                                                                                                      SOCTSCRGGLGGQCACAAAAA-------AAAAGCAGQGGYGGLGS 291
                                                                                                           AGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAA------AGGAGQGGYGGLG
                                                                                                                                                 GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                          GGQGAGAAAA----------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                    GGYGGLGSQGTSGRGGLGGQCAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA
                                                                                                                                                                                                                                                       agaaaaaaaaaaggggyggigsqgtsgrggiggqgagaaaaaaaaaagggggygg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5989894th Washington Street
CITY: Falls Church
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGVGSGASAA-----SAAASRLSSPOAS 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGLGSQGTSGPGGYGPGQQTSGIRIRAPSTS 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 2, Application US/08317844B Patent No. 5989894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: Virginia COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ZIP: 22046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-317-844B-2
                    82
                                                                                                202
                                                                                                                   143
                                                          142
                                                                                                                                      253
                                                                                                                                                                                                                                                                           371
                                                                                                                                                                                                                                                                                               453
                                                                                                                                                                                                                                                                                                                                   200
                                                                                                                                                         202
                                                                                                                                                                            292
                                                                                                                                                                                               261
                                                                                                                                                                                                                                                        393
                                                                                                                                                                                                                                                                                                                                                       176
                                                                                                                                                                                                                                                                                                                                                                          553
                                                                                                                                                                                                                                                                                                                                                                                                                 593
                                                                                                                                                                                                                                                                                                                                                                                                                                   288
                                                                                                                                                                                                                  351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         637
                                                         ò
                                                                           g
                                                                                               à
                                                                                                                 g
                                                                                                                                     ò
                                                                                                                                                        g
                                                                                                                                                                           ð
                                                                                                                                                                                           qq
                                                                                                                                                                                                                å
                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                    ð
                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
```

```
35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LGSQGTSGRGGLGGQGAGAAAAAAAGAGGGGGGGGGGGGT-----SGP 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120 LGNQG-AGRGGLGGQGAG---AAAAAAAGGAGQGGYGGLGNQGAGRGGQGAAAAAGGAGQ 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      371 AG---AVAAAAGGAGGGGYGGLGSQG----AGRGGQGAG----AAAAAAGGAGQRGYGG 419
                                                                                                                                                                                                                                                                                                                                                                                                             25 QGAGAAAAAAAAAAAGAGGGYGGLGSQGT--SGRGGLGGQGAG-AAAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                          1 QGAG----AAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGGQGAGQGAGAAAAAAAGGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OGTSGRGCLGGQGAGAAAAAAAAAAGAGGAGGGYGCLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGQGAGAAAA-------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         142 GAAAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SQGTSGRGGLGGQGAGAAAAAA........AAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGYGP-GQQTSGRGCLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     532 GGQGAGAAAAAGGAGGGLGGQGAGQGAGAAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   593 QGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA....-AGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                               Indels 184;
                                                                                                                                                                                                                                                                                                                                                  Length 718;
     #1.25
                                                                                                                                                                                                                                                                                                                                                                                  49:
                                                                                                                                                                                                                                                                                                                                             Ouery Match 60.3%; Score 2138; DB 2; Best Local Similarity 66.3%; Pred. No. 3.8e-147; Matches 498; Conservative 20; Mismatches 49;
SOFTWARE: Patentin Release #1.0, Version CURRENT APPLICATION NOMBER: US/08/317,844B FILING DATE: 04-OCT-1994
                                                              CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MULPHY Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/POCKET NUMBER: 1447-105P
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                             TELEX: 248345
INFORMATION FOR SEO ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 718 amino acids TYPE: amino acid
                                                                                                                                                                  TELEFHONE: (703) 241-1300
TELEFAX: (703) 241-2848
                                                                                                                                                                                                                                                                              ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-317-844B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGGGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    588
```

3

q à

```
QGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGAGGLGSQGTSGPGGYGP-GQQTSGRGGL 350
                                                                             ----- AAAAGGAGQGGYGGLGS 291
                                                                                                                                                                                                                                                                                                                GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                            QGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E: E. I. DU PONT DE NEMOURS AND COMPANY 1007 MARKET STREET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 19, Application US/08556978B
Patent No. 6268169
GENEAL INFORMATION:
APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
TITLE OF INVENTION: SPIDER SILK ANALOGS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
SOFTWARE: MICROSOFT WINDOWS 95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     653 GGLGSQGTSGPGGYGPGQQTSGIRIRAPSTS 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CR-9389-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: UNITED STATES OF AMERICA ZIP: 19898
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08/077,600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,60
FILING DATE: JUNE 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA ARAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9;
TELECPHONE: 302-892-8112
TELEPHONE: 302-773-0164
   SQGTSGRGGLGGQGAGAAAAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DELAWARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CITY: WILMINTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE:
STREET: 1(
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-556-978B-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE:
                                                                                                                                                       312
                                                                                                                                                                                                                                                                              420
                                                                                                                                                                                                                                                                                                              200
                                                                                                                                                                                                                                                                                                                                                                                                                                       593
                              202
                                                             292
                                                                                           261
                                                                                                                        351
                                                                                                                                                                                      393
                                                                                                                                                                                                                    37.1
                                                                                                                                                                                                                                                                                                                                             476
                                                                                                                                                                                                                                                                                                                                                                          553
                                                                                                                                                                                                                                                                                                                                                                                                         532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                637
                                                                                                                                                                                                                                                   453
                                                           ò
                                                                                        a
                                                                                                                        ð
                                                                                                                                                  g
                                                                                                                                                                                    ò
                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             82 GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAGGAGGGGYGGLGGSQGTSGRGGLGGQGA 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      108 G---AAAAAAGGAGGGYGGLGNQG------AGRGGQG------AAAAA 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25 QGAGAAAAAAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGAG-AAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60.3%; Score 2138; DB 3; Length 747; 66.3%; Pred. No. 3.9e-147; ive 20; Mismatches 49; Indels 18
                                                                                                  Sequence 3, Application US/09034177

Patent No. 6127146

GENERAL INFORMATION:
APPLICANT: Guegler, Karl J.
APPLICANT: Gorley, Neil C.
TITLE OF INVENTION: HUMAN FIBROUS PROTEIN
NUMBER OF SEQUENCES: 3
CORRESSONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/034,177
653 GGLGSQGTSGPGGYGPGQQTSGIRIRAPSTS 683
                           GGVGSGASAA-----SAAASRLSSPQAS 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: BILLINGS, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0486 US
TELECOMMUNICATION:
TELEPHONE: 650-855-0555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
LENGTH: 747 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE: HEREWITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: 650-845-4166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA: APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LIBRARY: GenBank
CLONE: GI 1174414
                                                                                                                                                                                                                                                             STREET: 3174 Port
CITY: Palo Alto
STATE: CA
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                             94304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Sim
Matches 498;
                                                                                        US-09-034-177-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-034-177-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                             637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   142
```

g ò qq 셤 ò

ò

AT

34;

```
419
          TITLE OF INVENTION: COSMETT OR DERMATOLOGICAL COMPOSITION CONTACTING A TITLE OF INVENTION: COSMETT OR DERMATOLOGICAL COMPOSITION CONTACTING TITLE OF INVENTION: ANALOG FILE REFERENCE: 6388-0365-0 CURRENT APPLICATION NUMBER: US/09/247,806 CURRENT APPLICATION NUMBER: US/09/247,806 CURRENT APPLICATION NUMBER: F98/01614 EARLIER APPLICATION NUMBER: F98/01614 SOFTHER FILING DATE: 1998-02-11 SARLIER FILING DATE: 1998-02-11 SOFTHARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                          350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82 GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 141
                                                                                                                                                                                                                                                                                            252
                                                                                                                                                                                                                                                                                                                                                                                                                   291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LGSQCTSGRGGLGGQGAGAAAAAAAAAAAGGGGGGGGGGGGGGGCT-----SGP 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         587
                                                                                                                                                                                                                                                         1 QGAG----AAAAAAGGAGQGGYGGLGGQGAGGGGAGAGAAAAAAGGAGQ 56
                                                                                                                                                                                                                Indels 176; Gaps
                                                                                                                                                                                                                                     25 QGAGAAAAAAAAAAGGAGGGGGGGGGGGGT--SGRGGLGGQGAG-AAAAAAAAAAAAGGAGQ 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA------AGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                  SQGTSGRGGLGGQGAGAAAAAA ------AAAAGGAGQGGYGGGS
                                                                                                                                                                                                                                                                                                                                                                                                                                QGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AGAAAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGYGP-GQQTSGRGGLGGQGAGAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGQGAGAAAA--------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGQGAGAAA----------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        532 GGGGAGAAAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                         Query Match 60.3%; Score 2135; DB 4;
Best Local Similarity 68.3%; Pred. No. 5.7e-147;
Matches 495; Conservative 15; Mismatches 39;
    APPLICANT: ARRAUDEAU, Jean-Pierre
                                                                                                                                                           ORGANISM: Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGVGS 641
                                                                                                                                     651
                                                                                                                                                                    US-09-247-806-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CCLGS
                                                                                                                                                 TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                393
                                                                                                                                                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                                                                                                                                                                                                                  253
                                                                                                                                                                                                                                                                                                                                                                                                                                       202
                                                                                                                                                                                                                                                                                                                                                                                                                                                             292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               637
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                          34;
                                                                                                                                                                                                                                                                                                                                               260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   350
                                                                                                                                                                                                                                                                                                                                                                                                                            GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 141
                                                                                                                                                                                                 G---AAAAAAAGGAGGGYGGLGNQG------AGRGGQG-----AAAAA 142
                                                                                                                                                                                                                                                                                252
                                                                                                                                                                                                                                                                                                   201
                                                                                                                                                                                                                                                                                                                          291
                                                                                                                                                                                                                                                                                                                                                                               392
                                                                                                                                                                                                                                                                                                                                                                                                                                                            452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGYGP-GQQTSGRGGLGGQGAGAAAAAAA-----AGGAQQGGYGGLGSQGTSGRGGL 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  587
                                                                                                                                                                                                                                   Gaps
                                                                                                                                                        25 QGAGAAAAAAAAAAGGAGGGGGGGGGGGGT--SGRGGLGGQGAG-AAAAAAAAAAAGGAGQ 81
                                                                                                                                                                                                                                                                                        AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAA...---AGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             532 GGGGAGAAAAAGGAGGGGGGGGGAGAAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SQGTSGRGGLGGQGAGAAAAAA-------AAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                    QGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                GGQGAGAAAA-------AAAAAAAGGAGGGGTGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QGAGAAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGY
                                                                                                                       39; Indels 176;
                                                                                                   Length 651;
                                                                                                60.3%; Score 2135; DB 4;
68.3%; Pred. No. 5.7e-147;
iive 15; Mismatches 39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1, Application US/09247806
Patent No. 6280747
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: PHILLIPPE, Michel APPLICANT: GARSON, Jean-Claude
                    LENGTH: 651 amino acids
                                                                                               Query Match
Best Local Similarity 68.3%
Matches 495; Conservative
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                TYPE: amino acid
STRANDEDNESS: unknown
                                                     ; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-556-978B-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGLGS 657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGVGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 5
US-09-247-806-1
                                                                                                                                                                                       82
                                                                                                                                                                                                            57
                                                                                                                                                                                                                                  142
                                                                                                                                                                                                                                                      108
                                                                                                                                                                                                                                                                            202
                                                                                                                                                                                                                                                                                                                                              202
                                                                                                                                                                                                                                                                                                                                                                   292
                                                                                                                                                                                                                                                                                                                                                                                         261
                                                                                                                                                                                                                                                                                                                                                                                                               351
                                                                                                                                                                                                                                                                                                                                                                                                                                     312
                                                                                                                                                                                                                                                                                                                                                                                                                                                          393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                653
                                                                                                                                                                                                                                                                                                                        253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      637
                                                                                                                                             ô
                                                                                                                                                                 g
                                                                                                                                                                                       ò
                                                                                                                                                                                                         g
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
```

201

260

```
Sequence 4, Application US/09247806

Patent No. 6280747

GENERAL INFORMATION:

APPLICANT: PHILLIPPE, Michel

APPLICANT: PHILLIPPE, Michel

APPLICANT: PRANUDEAU, Jean-Pierre

TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT

TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN

TITLE OF INVENTION: ANALOG

FILE REFERENCE: 6189 61056-0

CURRENT APPLICATION NUMBER: US/09/247,806

CURRENT FILING DATE: 1999-02-11

EARLIER FILING DATE: 1999-02-11

SAPERLER FILING DATE: 1998-02-11

NUMBER OF SEQ ID NOS: 14
AAGGACQCGYGGLGSQCT--SGRGGLGGQGA---GAAAAAAAAAAAAGGAGQGGYGGLGSQG 583
                                                                                                                                                                                                                81 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT--SGRGGLGG 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     55 Q------GLGSQGAGQGAGAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGS 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------GGLGS 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139 QGA----GAAAAAAAAAGGAGQGGYGGLGSQGT--SGPGGYGPGQQT-----SGRGGLG 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   243 AGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT--SGRGGLGGQGA---GAAAAAA
                                                                                                                              477 AAAAGGAGGGYGGLGSQGT--SGPGGYGPGQQT----SGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    58.6%; Score 2076; DB 4;
ilarity 69.7%; Pred. No. 9.5e-143;
Conservative 11; Mismatches 74;
                                                                                                                                                                                                                                                                                                                            644 AGGAGGGGYGGLGSQGTSGPGGYG 667
                                                                                                                                                                                                                                                                                                                                             Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Simi
Matches 477;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQ ID NO 4
LENGIH: 606
                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-247-806-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-247-806-4
                                                                                              359
                                                                                                                                                             417
                                                                                                                                                                                             529
                                                                                                                                                                                                                             477
                              엄
                                                              ò
                                                                                         g
                                                                                                                              ð
                                                                                                                                                           g
                                                                                                                                                                                           ò
                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                81 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT--SGRGGLGG 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  139 QGA---GAAAAAAAAAAAGGAGGGGGGGGGGGGT--SGPGGYGPGQQT----SGRGGLG 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 243 AGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      188 GQGAGAAAAAAAAAAGGAGGGGGGGGGGGT---SGRGGLGGQGA---GAAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74; Indels 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 606;
                                                                       GENERAL INFORMATION:
APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED TITLE OF INVENTION: SPIDER SILK ANALOGS NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: I. DU PONT DE NEMOURS AND COMPANY STREET: 1007 MARKET STREET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 69.7%; Pred. No. 9.5e-143;
Matches 477; Conservative 11; Mismatches 74;
                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9389-A
TELECOMMUNICATION:
TELEPHONE: 302-8912
                                                                                                                                                                                                                                         UNITED STATES OF AMERICA
                                            Sequence 21, Application US/08556978B Patent No. 6268169
                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,600
FILLING DATE: JUNE 15, 1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 21:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: unknown
MOLECULE TYPE: protein
                                                                                                                                                                                        STREET: 1007 MARK
CITY: WILMINTON
STATE: DELAWARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRANDEDNESS:
                                                                                                                                                                                                                                       COUNTRY: UI
                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-556-978B-21
                                                            Patent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

30;

à

ŏ 8 ò à

9

303	QGAGAAAAAAAAAGGAGGGGGGGGGGGGGGGGGGGGGG		Query Match Best Local	Query Match Best Local Simila
Db 263 QGAGQGAGAAAAAGGAGQGGYGGLGSQG-AGQGGYGGLGSQGAGRGGQGAG	G-AGQGGYGGLGSQGAGRGGQGAGA 314		Matches	s 462; Cor
Qy 362 AAAAAGGAGQGGYGGLGSQGTSGRGGLG	AAAAAGGAGGGGYGGLGSQGTSGRGGIGGQGAGAAAAAAAAGGAGGGGGGGGGGGGGGGGGG	VO OY		40 GAGOGGYGG
Db 315 AAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG		qa		1 GAGOGGYGG
Oy 422 GRGGLGCQCAGAAAAAAAGGAGQGG	GRGGLGGQGAGAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAA 476	ον	100	g
Db 359GGLGSQGAGQGAGAAAAAGGAGQGG	GELGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGGGGGGGAAAAAAAAAAAAAAAAAAAAAAAAA	qO	43	SQGAGQGAG
Qy 477 AAAAGGAGGGYGGLGSQGTSGPGGYGPGQQT	GPGQQTSGRGGLGGQGAGAAAAAAA 528	δ 	160	yegresogr:
Db 417 AAAAGGAGQGGYGGLGSQGAGRGGLGG		q _Q	95	YGGLGSQG-1
Qy 529 AAGGAGGGGGGGGGGGTSGRGGLGGGGA	AAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGGGGGGGGGGGGGG 583	Oy.	198	AAAAAGGAG
Db 477 AAGGAGQGGYGGLGSQGAGQGGYGGLGS	GRGGQG	qa	154	AAAAAGGAG
Oy 584 TSGRGGLGGGGAGAAAAAAGGAG	TSGRGGLGGGGGGAAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGGGAGAAAAAAA 643	ογ	257	
	GGTGSÖGAGÖG	qa	206	S QGGYGGLGSQ
Oy 644 AGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG		δ	306	GAAAA
Db 579 AGGAGGGYGGLGSQG-AGGGGYG 601		q	249	GOGAGAAAA
RECITI'A B		νο	366	-
US-08-556-978B-23	а	qa	289	AGGAGQGGYC
Partent No. 6268169	1	δδ	414	GLĠSQGTSGE
APPLICANT FAINESTOCK, STEPHEN F	THE PERSON NAME OF THE PERSON NA	qa	342	
TILL OF INVENTION: NOVEL RECOMMENTI FRODUCED TILL OF INVENTION: SPIDER SILK ANALOGS NUMBER OF SECURIORES: 107	INANILI FRODUCED ANALOGS	0y	474	AAAAAAAGG
CORRESPONDENCE ADDRESS:	MOTTES AND COMBANY	qa	385	- AAAAAAGG
$^{\circ}$	MOOKS AND COMPANI	, 0	525	
STATE: DELAMINED COUNTRY. HANTED CHARES OF AMERICA	V C	q ₀	443	GGLGSQGAGC
ZIPS 3 ZIPS CHILL CHILD OF CALLS		<i>λ</i> ο	572	GOGGYGGLGS
7 2	CH	qo	495	
; COMPUTER: IBM PC COMPATIBLE ; OPERATING SYSTEM: MICROSOFT WI	NDOWS 95	^0	601	AAAAAAGG
SOFTWARE: MICROSOFT WORD FOR WINDOWS CURRENT APPLICATION DATA:	INDOWS 95			
APPLICATION NUMBER: US/08/556,978B	978в	<u> </u>		AGAGAGAGG
CLASSIFICATION: 435		RE	RESULT 9	
; PRIOR APPLICATION DATA: ; APPLICATION NUMBER: 08/077,600		SD .	-09-247 Sequence	US-09-247-806-8
; FILING DATE: JUNE 15, 1993 : ATTORNEY/AGENT INFORMATION:			Patent	Patent No. 6280747
NAME: FLOYD, LINDA AXAMETHY			APPLIC	ANT: PHILL
REFERENCE/DOCKET NUMBER: CR-9389-A	89-A		APPLIC APPLIC	APPLICANT: GARSON APPLICANT: ARRAUI
; TELECOMMUNICATION INFORMATION: ; TELEPHONE: 302-892-8112			TITLE	OF INVENTI
; TELEFAX: 302-773-0164			TITLE	OF INVENTI
; INFORMATION FOR SEQ ID NO: 23: ; SEQUENCE CHARACTERISTICS:			FILE F	FILE REFERENCE: CURRENT APPLICAT
; LENGTH: 606 amino acids ; TYPE: amino acid			CURRE	TFILING D
S			EARLIE	SR FILLING D
MOLECULE TYPE: protein			NUMBER	NUMBER OF SEQ ID SOFTWARE: Patent1
US-U8-956-87-88-Z3			SEQ ID NO 8	NO 8 TH: 606
			i	

```
29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HILLIPPE, Michel
ARSON, Jean-Claude
ARSON, Jean-Claude
ARSON, Jean-Pierre
RRADDEAU, Jean-Pierre
ENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
ENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
ENTION: LANALOG
CE: 6388-0365-0
ICATION NUMBER: US/09/247,806
NG DATE: 1999-02-11
ICATION NUMBER: FR 98/01614
OLD NOS: 14
Cold NOS: 14
Cold NOS: 14
Cold NOS: 14
Cold NOS: 14
                                                                                                                                                                                   305
                                                                                                                                                                                                                                                                                                                                                                            473
                                                                                                                                                  YGGLGSQGT------SGRGGLGGQGAGAAAAAAAAGAGGGGGGG 413
|||||||||
YGGLGSQGAGQGGYGGLGSQGAGGGGGGGGG----341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  rscpggygp-gootsgrgglgggagaaa------- 197
                                                                                                                                                                                                                                                                                                                                                                                       --GGLGSQGAGQGAGAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG--- 384
                                                                                                                                                                                                                                                                                                                                                                                                                         524
                                                                                                                                                                                                                                                                                                                                                                                                                                              442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        554
                        Indels 215; Gaps
                                                          94
                                              ---SGRGGLGGQGAGAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                     3RGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA
55.7%; Score 1972.5; DB 4; Length 606; 64.2%; Pred. No. 2.7e-135; i.ve 8; Mismatches 35; Indels 215;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ication US/09247806
47
                       nservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 606
```

```
GGLGSQGT-----SGRGGLGGQGAGAAAA------AAAA
                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
                                                                                                                                                                                                                                                  STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                             TOPOLOGY: unknown
                                                                                                                                                                                                                                        TYPE: amino acid
                                                                                                                                                                                                                                                                                   US-08-556-978B-63
                                                                                                                                                                                                                                                                                                                                                                                                                                                          97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          257
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                               29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             206 QGGYGGLGSQGAGRGGLGGQGAG---AAAAAAAGGAGQ--------GGLGSQGA 248
                                                                                                                                                                                                                                                                                                                  524
                                                                                                                                                                                                                                                                                                                                                                                                                                                          442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----SGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 305
                                                                                                                                                                                                                                                                                                      306 GAAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAA 365
                                                                                                                                                                                                                                                                                                                                                366 AGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAAAGGGGGGYG 413
                                                                                                                                                                                                                                                                                                                                                                      341
                                                                                                                                                                                                                                                                                                                                                                                           GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA 473
                                                                                                                                                                                                                                                                                                                                                                                                       YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-------A 197
                                                               Gaps
                                                                                             GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 99
                                                                                                                                                  SQGAGQCAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AAAAAAGGAGQGG 94
                                                                                                                                                                                                                                                                                                                                                                                                                                  AAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGYGGLGSQGT~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            495 GOGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGQGGLGSQGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --AAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGGYGGLGSQ
                                        Length 606;
                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GQGGYGGLGSQGT----SGRGGLGGQGAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
TITLE OF INVENTION: SPIDER SILK ANALOGS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
                                          DB 4;
                                       Score 1972.5; DB 4
Pred. No. 2.7e-135;
8; Mismatches 35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-556-978B-63; Sequence 63, Application US/08556978B; Patent No. 6268169
                                      55.7%;
illarity 64.2%;
Conservative 8
       ; ORGANISM: Nephila clavipes US-09-247-806-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CITY: WILMINTON
STATE: DELAWARE
COUNTRY: UNITED &
                                                  Similarity
                                                 Best Local Similarity Matches 462;
                                         Query Match
                                                                                  40
                                                                                                                                                  43
                                                                                                                                                                                                                198
                                                                                                                                                                                                                                                           257
                                                                                                                                                                                                                                                                                                                                                                                           414
                                                                                                                                                                                                                                                                                                                                                                                                               342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   601
                                                                                                    g
                                                                                                                                                 g
                                                                                                                                                                                      g
                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                             à
                                                                                                                                                                        ò
                                                                                                                                                                                                                δý
                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        . . . . . . . . . . . . . . . .
```

```
305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        160 YGGLGSQGTSGPGGYGP-GQQTSGRGCLGGQGAGAAA------A 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100 GQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAAAGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      198 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT•
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----SGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           208 QCCYCCLCSQCACRCCLCCQCAG---AAAAAAGCACQ-------CGLCSQCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         251 GOGAGAAAAAAGGAGGGYGGLGSQG-----AGRGGQGAGAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              479 AAGGAGQGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----AAAAAGGAGGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAGGAGGGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 604;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 1971.5; DB 4;
Pred. No. 3.1e-135;
8; Mismatches 37;
        COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
                                                  SOFTWARE: MICROSOFT WORD FOR WINDOWS
CURRENT APPLICATION DATA:
APPLICATION NUBER: US/08/556,978B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: 08/077,600
FILING DATE: JUNE 15, 1993
ATTOMAYAGENT IRFORMATION:
NAME: FLOYD, LINDA AXAMETHY
                                                                                                                                                                                                                                    NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9389-A
MEDIUM TYPE: DISKETTE, 3.50 INCH
                                                                                                                                                                                                                                                                                               TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELERAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 63:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 55.6%;
Best Local Similarity 64.7%;
Matches 461; Conservative 6
```

29;

œ

Page

```
79 GGYGAG---AGAVAAAGAAGAGGYG----RGAGGYGGQGGYGAGAGAAAAA--GAGAG 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYG----GLGSQGTSGRGGLG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 124; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        601 --AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLG 656
                                                                                                                                                      Length 832;
                                                                                                                                                                                                                                                                                                               APPLICANT: CALGIN, Mark
APPLICANT: Colgin, Mark
TITLE OF INVENTION: CDNAs Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSE: Birch, Stewart, Kolasch & Birch
STREET: P.O. BOX 747
CITY: Falls Church
                                                         572 GQGGYGGLGSQGT-----SGRGGLGGQGAGAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,747
FILING DATE: 14-MAR-1994
CLASSIFICATION: 530
ATTONEY/AGENT INFORMATION:
NAME: MULPHY Jr. Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 39.7%; Score 1405; DB 1; 18est Local Similarity 50.1%; Pred. No. 2.4e-94; Matches 349; Conservative 17; Mismatches 207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM: N. clavipes
TISSUE TYPE: minor ampullate gland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                             Sequence 2, Application US/08209747; Patent No. 5733771; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELERAX: 703-205-8030
INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 832 amino acids TYPE: amino acid STRANDEDNESS: single TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1..309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: pr
HYPOTHETICAL: NO
FRAGMENT TYPE: in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; LOCATION:
US-08-209-747-2
                                                                                                                                                                                                                            RESULT 12
US-08-209-747-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURE
                                                                                            g
                                                                                                                               ŏ
                                                                                                                                                                  셤
                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29;
                                                                                                                                                                                                                                  APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: ARRANDEAU, Jean-Plerre
TITLE OF INVENTION: COSNETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: AMALOG
FILE REFERENCE: 6386-0365-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER PELING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
208 QGGYGGLGSQGAGRGGLGGGGAGG---AAAAAAAGGAGQ-------GGLGSQGA 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            160 YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA---------A 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYG 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               414 GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    344 ------GELGSQGAGQGAGAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG--- 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGA 571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     40 GAGOGGYGGLGSOGTSGRGGLGGOGAGAAAAAAAGGAGOGGYGGLGSOGTSGRGGLG 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 GAGGGGYGGLGSQG-AGRGGLGGQGAG---AAAAAAAGGAGQ------GGLG 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            474 AAAAAAAGGAGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      35; Indels 215;
                                                      604 AAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLG 656
                                                                           Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 1961.5; DB 4
Pred. No. 1.6e-134;
8; Mismatches 35;
                                                                                                                                                                               Sequence 6, Application US/09247806
Patent No. 6280747
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 64.1%;
Matches 460; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT ORGANISM: Nephila clavipes US-09-247-806-6
                                                                                                                                           RESULT 11
US-09-247-806-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                525
                                                                                     셤
                 g
                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

26;

σ

```
83
                                                                                                                                                                                                                                                                                                                                                                                                         138
                                                                                                                                                                                                                                                                                                                                                                                                                                                         198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                               470
                                                                                                                                                                                                                                             450
                                                                                                                                                                                                                                                                    521
                                                                                                                                                                                                                                                                                            505
                                                                                                                                                                                                                                                                                                                     577
                                                                                                                                                                                                                                                                                                                                             556
                                                                                                                                                                                                                                                                                                                                                                     637
GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAA 197
            -- AAAAAG
                                                                    242 GAAAGAGAGAGSYG-----GQGYGAGAGAAAAXXXXXXXXXXXXXXXXGAGAGA
                                                                                              242 GAGQGGYG----GLGSQGTSGRGGLGGQGAQAAAAAAAAAAAGGGGGYGG-----
                                                                                                                                             ----LGSQGTSGRGGLG---GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPG
                                                                                                                                                                                           342 QQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQG------TSGRGGLGG
                                                                                                                                                                                                                   413 AAGAGGAGGYGRGAGAGAGAAA--GAGAGGYGCQGGYGAGAGAGAAAAATGAGGAGG
                                                                                                                                                                                                                                            451 G-GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG---LGSQGTSGPG-GYGPG
                                                                                                                                                                                                                                                                                                                                            506 QQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYG---GLGS----QGTSGRGGLG-GQG
                                                                                                                                                                                                                                                                                                                                                                                          557 AGAAAAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 2, Application US/08458298

Patent No. 575677

GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Lewis, Mark
TITLE OF INVENTION: SINk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STREET: Viginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM FC COMPATIBLE
OPERATUG SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                          617 LGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYG 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                        - RGAGAGGYGGGGYGAGAGAAAA--GAGAGGYG 717
                                             198 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,747
FILING DATE: 14-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: MUTPHY JI., GETALD M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: US/08/458,298
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: USA
ZIP: 22040-3487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-458-298-2
ò
                    q
                                                                   셤
                                                                                                                셤
                                                                                                                                       ò
                                                                                                                                                                qq
                                                                                                                                                                                                             a
                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                             ò
                                                                                             ò
                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             151 G-GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGGGGGGGGG---LGSQGTSGPG-GYGPG 505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       506 QQTSGRGCLGGQGAGAAAAAAAAAGGAGGGGYG---GLGS----QGTSGRGGLG-GQG 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGYGAG---AGAVAAAGAAGAGGYG----RGAGGYGGQGGYGAGAGAAAA--GAGAG 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GYGGLGSQGTSGRGGLGGQGAGAAAAAAAGGAGQGGYG----GLGSQGTSGRGGLG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GQGAGAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAA 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAAAAA 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----LGSQGTSGRGGLG---GQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPG 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19.7%; Score 1405; DB 1; Length 832;
Similarity 50.1%; Pred. No. 2.4e-94;
Conservative 17; Mismatches 207; Indels 124; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23 GGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQG 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    342 QQTSGRGGLGGQGAGAAAAAAAAGGGGGGGGGCGCGCGC-----TSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         242 GAGQGGYG-----GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYG 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: N. clavipes
TISSUE TYPE: minor ampullate gland
                     1447-104P
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 14,
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEPAX: 703-205-8050
                                                                     INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 832 anino acids TYPE: amino acid STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                LOCATION:
                                                                                                                                                                                                                                                                                                                                              NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 349;
                                                                                                                                                                                                                                                                                                                                                                                US-08-458-298-2
                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local
```

```
Sequence 10, Application US/09247806

Patent No. 6280747

GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARNOLEGA, Jean-Claude
APPLICANT: GARNOLEGA, Jean-Pierre
TITLE OF INVENTION: COSMETIC ON DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: LAST ONE NATURAL OR SECREDIAL CONTACTING AT
TITLE OF INVENTION: LAST ONE NATURAL OR SECREDIAL CONTACTING AN
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER APPLICATION UNMER: FR 98/01614
EARLIER PILING DATE: 1998-02-11
CONTACTION OF THE ORDER ORDE
                                                                                                                                                              407 AAAAGPGGYGP-GQQGPGGYGP-GQQGPSGPAAAAAAAAAAGPGGYGP-GQQGPGGYGP 463
                                                                                                                                                                                                                                                                                                                                                                          524 AAAAAAGPGGYGPGQGPGGPGGPSGPSGAAAAAAAAAAAGPGGYGP-GQQGPGGYGP 582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RGGLGGQGAGAAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG-GQCACAAAAAA 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            171 PGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG- 225
                          SGRGGLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAA
                                                                                                           AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRG-
                                                                                                                                                                                                                             ---GLGGQGAG-----AAAAAAAAAAAGGAGQGGYG------GLGSQGTS----
                                                                                                                                                                                                                                                                                   464 GQQGPGGYGPGQQGPSGPGSAAAAAAGPGQQGPGGYGPGGYGPGQQGPSGPGSAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---GLGGQGAG----AAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLG-GQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    583 GOGGPGGYGPGQQGPSGPGSAAAAAAGPGQQGPGGYG-PGQQGPGGYGPGQQGPSGPGSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAAAGGAGQGGYG-GLGSQGTSGRGGLGGQGAGAAAAAAAAAAG----GAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 ASMTGGQQMGRGSMASGRGGLGGQGAG----AAAAAAAAAAAAGGGGGYG-GLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAGGAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----GPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGGGGYGGLGSQGTSGRG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              38.0%; Score 1346.5; DB 4; Length 714; 50.0%; Pred. No. 3.4e-90; Live 26; Mismatches 248; Indels 93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; OTHER INFORMATION: Description of Artificial Sequence:protein US-09-247-806-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               658 QGTSGPGGYGPGQQ 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best_Local Similarity
Matches 367; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-247-806-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                             463
                                                                                                                                                                                                                                                                                                                                               498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           57
                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAAGGAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSG 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG- 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PGOOGPSGPGSAAAAAAAAGPGGYGPGQYGPGGYG-PGQQGPGGYGPGQQGPSGPGSAA 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          57 RGGLGGQGAGAAAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAAA 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10 AAAAGPGQGGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYGPGQQGPGGYG 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 ASMTGGQQMGRGSMASGRGGLGGQGAG----AAAAAAAAAAGGAGQGGYG-GLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              38.0%; Score 1346.5; DB 4; Length 714; 50.0%; Pred. No. 3.4e-90; Live 26; Mismatches 248; Indels 93;
                                                                                                                                                                                                                                                                                                        : E. I. DU PONT DE NEMOURS AND COMPANY 1007 MARKET STREET
                                                                                                                                                                                TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED TITLE OF INVENTION: SPIDER SILK ANALOGS NUMBER OF SEQUENCES: 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CR-9389-A
                                                                                                                                                                                                                                                                                                                                                                                                                          UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,600
FILING DATE: JUNE 15, 1993
ATTONNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETH
REGISTRATION NUMBER: 33,692
                                                                 Sequence 61, Application US/08556978B Patent No. 6568169
GENERAL INFORMATION:
GAPPLICANT: FAHNESTOCK, STEPHEN F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE/DOCKET NUMBER: CR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO: 61:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
LENGTH: 714 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 50.09
Matches 367; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    302-773-0164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: unknown
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                 CITY: WILMINTON STATE: DELAWARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRANDEDNESS:
                                                                                                                                                                                                                                                                                                           ADDRESSEE:
           RESULT 14
US-08-556-978B-61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-556-978B-61
                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

35;

Gaps

```
        603
        AAAAGGAGGGYG-GLGSOGTSGRGGLGGGGAAAAAAAAAG----CAGGGGYGGLGS
        657

        1111
        1111
        111
        111

        642
        AAAAAAAGGGYGFGQGFGGYGFGQGFSGPGSAAAAAAAAAAAGGGYGFGQGFGGYG-700
        700

GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQT 344
                                                           497
                                                                                                                                                  284
                                                                                                                                         ---GLGGQGAG-----AAAAAAAAAAAGAGGGYG------GLGSQGTS----
                     226 GQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                               658 QGTSGPGGYGPGQQ 671
                                                                                                                                                                                                                                                                       285
                                                                                                                                          463
                                 g
                                                 g
                                                                                                                                                                                                 g ç
                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                              ò
```

Search completed: December 18, 2002, 16:23:09 Job time: 18:9094 secs

This Page Blank (uspto)

```
Sequence 1, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 10, Appli
Sequence 27, Appli
Sequence 27, Appli
Sequence 269, Appli
Sequence 27, Appli
Sequence 27, Appli
Sequence 27, Appli
Sequence 27, Appli
Sequence 37, Appli
Sequence 525, Appli
Sequence 451, Appli
                                                                                                                                           December 18, 2002, 16:21:10 ; Search time 9.54566 Seconds (without alignments) 1213.493 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                US-09-490-291-6
3543
1 MASWIGGQWGRGSMASGRG......TSGIRIRAPSISFEHHHHHH 691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /cgn2_6/ptodata/1/pubpaa/USO8_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO6_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO6_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO7_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO7_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO8_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO8_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO8_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO8_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO8_PUBCOMB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-10-029-217A-24
US-10-096-961-1
US-09-879-957-37
US-09-756-0718-20
US-09-756-0718-20
US-09-756-0718-20
US-09-735-931
US-09-735-933-1
US-09-735-933-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-861-597-1
US-09-861-597-4
US-09-861-597-8
US-09-861-597-6
US-09-861-597-2
US-10-052-586-97
US-10-052-586-97
US-10-052-586-26
US-10-052-586-26
US-10-052-586-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                      106657 seqs, 16763532 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                   OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0070700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          55580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580

5580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1972.5
1961.5
1346.5
1098.5
863.5
742
687.5
675.5
                                                                                                                                                                                                                                                                              Perfect score:
                                                                                                                                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      652
651.5
610
                                                                                                                                                                                                                                                                                                           Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                      Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                                                                                                   Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
```

	Sequence 23, Appl Sequence 36, Appl Sequence 25, Appl
US-09-975-143-47 US-10-052-586-581 US-09-19-497-56 US-09-864-761-36720 US-09-841-321A-37 US-09-841-321A-37 US-09-841-321A-34 US-09-837-969A-34 US-09-837-969A-39 US-09-837-969A-39 US-09-837-969A-39 US-09-841-321A-38 US-09-841-321A-38 US-09-841-321A-38 US-09-841-321A-38 US-09-841-321A-38 US-09-967-624-18 US-09-967-624-18 US-09-841-321A-18 US-09-841-321A-18 US-09-841-321A-18 US-09-841-321A-18 US-09-841-321A-18	US-09-83/-909A-25 US-09-837-969A-36 US-09-841-321A-25
	200
914 4 1 1 4 3 5 1 1 8 1 8 1 8 1 8 1 8 1 8 1 8 1 8 1 8	635 635 635 635 635 635 635 635 635 635
11.2.2.3.3.4.4.5.5.6.4.4.8.8.3.3.2.2.6.8.3.3.3.3.4.8.3.5.6.6.4.4.8.8.3.3.3.3.3.3.3.3.3.3.3.3.3.3.3.3	11.8
44 00 4444	419.5 419.5 419.5
01000000000000000000000000000000000000	4 4 4 U 4 10

ALIGNMENTS

Sequence 1, Application US/09861597

Patent No. US20020064539A1

GENERAL INFORMATION:

APPLICANT: PHILLIPPE, Michel

APPLICANT: GARSON, Jean-Claude

APPLICANT: GARSON, Jean-Pierre

TITLE OF INVEWTION: COSMETTO R DERNATOLOGICAL COMPOSITION CONTACTING AT

TITLE OF INVEWTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN

TITLE OF INVEWTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN

FILE REFERENCE: 6388-0365-0

CURRENT APPLICATION NUMBER: US/09/861,597

CURRENT APPLICATION NUMBER: 09/247,806

PRIOR FILING DATE: 1999-02-11

PRIOR FILING DATE: 1999-02-11

NUMBER OF SEQ ID NOS: 14

SEQ ID NO 1

LENGTH: 651 34; Indels 176; Gaps Length 651; Query Match 60.3%; Score 2135; DB 10; Best Local Similarity 68.3%; Pred. No. 2.8e-120; Matches 495; Conservative 15; Mismatches 39; TYPE: PRT ORGANISM: Nephila clavipes US-09-861-597-1 US-09-861-597-1 g g 셤 à ð ò

```
3 GRGGQGAG----AAAAAAGGAGQGGYGGLGSQG-AGRGGLGGQGAG----AAAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
US-09-861-597-8
                                                                                                                                                                                                                               263
                                                                                                                                                                                                                                                                                                                                                                                       529
                                                                                                                                                                                                                                                                                                                                            477
                                                                                                                                                                                                                                                                                                                                                                                                            q
                          ò
                                              셤
                                                                     ò
                                                                                        g
                                                                                                                 ð
                                                                                                                                      g
                                                                                                                                                            ò
                                                                                                                                                                               셤
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                  원
                                                                                                                                                                                                                                                                                                                                           å
                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: PHILLIPPE, Michel
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Plerre
TITLE OF INVENTION: CEMESTIC OR DERHATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6380-0365-0
CURRENT APPLICATION NUMBER: US/09/861,597
PRIOR APPLICATION NUMBER: 09/247,806
PRIOR PILLING DATE: 1999-02-11
PRIOR FILING DATE: 1999-02-11
PRIOR FILING DATE: 1999-02-11
PRIOR FILING DATE: 1999-02-11
SPRIOR FILING DATE: 1999-02-11
SPRIOR FILING DATE: 1999-02-11
SPRIOR FILING DATE: 1999-02-11
SPRIOR FILING DATE: 1999-02-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30;
                                                                                        312 GGQGAGAAAAGGAGGGLGGQGAGQAGAAAAAAGGAGGGYGGLGSQG-AGRGGLGGQG 370
                                                                                                                                                                                                                                                                                                                   GGQGAGAAAA---------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 592
                                                                                                                                                                                                                                                                                                                                                                 652
                                                                                                                                                                                                                                                                                                                                                                             201
                                              291
                                                                                                                                                                                 LGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGGGGGGCGGGGGGGT-----SGP 499
                                                                                                                                                                                                                                                                        GGYGP-GQQTSGRGGLGGQGAGAAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL 552
                                                                                                                                                                                                                                                                                    587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLGGGGAGAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAG 80
532 GCQGAGAAAAAGGAGQGGLGGQGAGQGAGAAAAAGGVRQGGYGGLGSQG----AGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 58.6%; Score 2076; DB 10 Best Local Similarity 69.7%; Pred. No. 8.2e-117 Matches 477; Conservative 11; Mismatches 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 4, Application US/09861597
Patent No. US20020064539A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; ORGANISM: Nephila clavipes US-09-861-597-4
                                                                                                                                                                                                                                                                                                                                                                                                                      ||:||
|GVGS 641
                                                                                                                                                                                                                                                                                                                                                                                                          653 GGLGS 657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 606
TYPE: PRT
                                                                                                                                                                                                                            453
                                                                                                                                                                                 393
                                                                   202
                                                                                           292
                                                                                                               261
                                                                                                                                                                                                                                                                        200
                                                                                                                                                                                                                                                                                             176
                                                                                                                                                                                                                                                                                                                   553
                                                                                                                                                                                                                                                                                                                                                                593
                                                                                                                                     351
                                                                                                                                                                                                                                                                                                                                                                                                                                637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21
                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                  8
  ô
                      셤
                                                                 ద
                                                                                        õ
                                                                                                              g
                                                                                                                                     à
                                                                                                                                                                                à
                                                                                                                                                                                                     g
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                              ç
```

```
Sequence 8, Application US/09861597

Fatent No. US20020064539A1

GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: PHILLIPPE, Michel
APPLICANT: PHILLIPPE, Michel
APPLICANT: PHILLIPPE, Michel
APPLICANT: PRRAUDEAU, Jean-Pierre
APPLICANT: PRRAUDEAU, Jean-Pierre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
ITILE OF INVENTION: ANALOG
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT PILLING DATE: 1999-02-11
PRIOR PELICATION NUMBER: ER 98/01614
PRIOR PELING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 8
LENGTH: 606
                                                                                                                                         160
                                                                                                                                                                           242
                                                                                                                                                                                                                                                                                                                                                                                                                                                        528
                                                                   100
                                                                                                      187
                                                                                                                                                                                                          161 SQGAGQGAGAAAAAAGGAGGGYGGLGSQGAGQGGYGGLGSQGAGRGGQGAGAAAAAGG 220
                                                                                                                                                                                                                                                               417 AAAAGGAGGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGGGGLGSQGAGQGAGAAAA 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         303 QGAGAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAAA 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    584 TSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 643
                                                                                                                                                                                                                                                                                                                                                                                    54
                               QGGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGT--SGRGGLGG
                                                              139 QGA---GAAAAAAAAAAGGAGGGYGGLGSQGT---SGPGGYGPGQQT----SGRGGLG
                                                                                                                                                                        188 GQGAGAAAAAAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGA---GAAAAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   644 AGGAGQGGYGGLGSQGTSGPGGYG 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; TYPE: PRT; ORGANISM: Nephila clavipes
US-09-861-597-8
```

m

29;

us-09-490-291-6.rapb

```
Sequence 10, Application US/09861597
Patent No. US20020064539A1
GENERAL INFORMATION:
APPLICANT: HILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Pierre
TILLE OF INVENTION: LEAST ONE DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: MAALOG
FILE REPERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT FILING DATE: 2001-05-22
                                                                                                                                                                                                                                                                                                                                                                                                                               290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQG----AGRGGQGAG----AAAAAAAGGA 496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        900
                                                                                                                                                                                    250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           197
                                                                                                                                                                                                                                                                                                            305
                                                                                                                                                                                                                                                                                                                                                                                                      AGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAAGGGGGGGG 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  344 -----GGLGSQGAGQGAGAAAAAGGAGGGGYGGLGSQG----AGRGGQGAG--- 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                3 GAGQGGXGGLGSQG-AGRGGLGGQGAG---AAAAAAAGGAGQ-------GGLG 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          601 --AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGGYGGLG 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                 QGGYGGLGSQGAGRGGLGGQGAG--------GGLGSQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GQGGYGGLGSQGT-----SGRGGLGGQGAGAAA-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  160 YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAAA------A
                                                                                                                                                                                                                                                                                                                                                   AAAAAGGAGQGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAAAAAAAGGAGQGGYGGLGSQGT-
                                                                                                        Indels 215;
                                                                                Length 606;
                                                                             Score 1961.5; DB 10;
Pred. No. 4.9e-110;
8; Mismatches 35; In
                                                                              Query Match
Best Local Similarity 64.1%;
Matches 460; Conservative
                                  ; ORGANISM: Nephila clavipes US-09-861-597-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 5
US-09-861-597-10
 SEQ ID NO 6
LENGTH: 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                        366
                                                                                                                                                                                                                                                                                             198
                                                                                                                                                                                                                                                                                                                                                257
                                                                                                                                                                                                                                                                                                                                                                           208
                                                                                                                                                                                                                                                                                                                                                                                                     306
                                                                                                                                                                                                                                                                                                                                                                                                                              251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                          a
                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                       à
                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SERVEMENT INCOME.

APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0361, 597
CURRENT FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US/09/47,806
PRIOR FILING DATE: 1999-02-11
PRIOR FILING DATE: 1999-02-11
PRIOR PILING DATE: 1998-02-11
SPIOR APPLICATION NUMBER: FR 98/01614
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver: 2.1
                           29;
                                                                                                                                                                                                                                                                   305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           442
                                                                                                                                                                                                                256
                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                    95 YGGLGSQG-AGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAGA 153
                                                                                                                                                                                                                              206 QGGYGGLGSQGAGRGGLGGQGAG---AAAAAAAGGAGQ-------GGLGSQGA 248
                                                                                                                                                                                                                                                                                                                       366 AGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYG 413
                                                                                                                                                                                                                                                                                                                                                                                                     341
                                                                                                                                                                                                                                                                                                                                                                                                                                                        384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  554
                                                                                                       .----- 197
                                                                                                                                                                                                                                                                                                                                                                                                                            414 GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------AAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGA 571
                           Gaps
                                                                   94
                                                   40 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 99
                                                                                                                   289 AGGAGQGGYGGLGSQGAGGGGGGGGGGGGGGGGGGGGG---AAAAAAAGGAGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              495 GOGGYGGLGSQGAGGGGGGGGGGGGGGGGAGAAAAAAGGAGGGGLGSQGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                          AAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GQGGYGGLGSQGT-----SGRGGLGGQGAGAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 606
                           35; Indels 215;
                                                                                                                                                          160 YGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAA-----
 DB 10;
Score 1972.5; DB 10
Pred. No. 1.1e-110;
8; Mismatches 35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 6, Application US/09861597 Patent No. US20020064539A1 GENERAL INFORMATION:
55.78;
64.28;
                           462; Conservative
             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-861-597-6
 Query Match
                           Matches
                                                                                                       100
                                                                                                                                                                                                               198
                                                                                                                                                                                                                                                                                                                      306
                                                                                                                                                                                                                                                                                                                                                                                                                                                      342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          601
                                                                                                                                                                                                                                                                   257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                    ð
                                                                             g
                                                                                                       ò
                                                                                                                             셤
                                                                                                                                                          å
                                                                                                                                                                                  셤
                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                  35;
                                                                                                                                                                                                                                                                                                                                                 70 PGQQGPSGPGSAAAAAAAAAAAGPGGYGPGQQGPGGYG-PGQQGPGGYGPGQQGPSGPGSAA 128
                                                                                                                                                                                                                                                                                                                                                                            PGGYGPGQQ-----GPSGPGSAAAAAAAAAAAGPGGYGPGQQGPGGYG-PGQQGPGGYGP 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GOOGPSGPGSAAAAAAGPGQGPGGYG-PGQGPGGYGPGQGPSGPGSAAAAAAAAGPG 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAA 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PGGYGPGQYGPSGPGSAAAAA-GPGQQGPGGYG-PGQQGPGGYGPGQQGPSGPGSAAAA 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG- 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       464 GQQGPGGYGPGQQGPSGPGSAAAAAAGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAA 523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              524 AAAAAAGPGGYGPGQYGPGQQGPSGPGSAAAAAAAAAAGPGGYGP-GQQGPGGYGP 582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RGGLGCQGAGAAAAAAAAAG----GACQGGYGGLGSQGTSGRGGLG-GQGAGAAAAAA 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLG-GQGAGAAAAAAAAAAGGAGQG 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      583 GOGGPGGYGPGOOGPSGPGSAAAAAAAGPGOOGPGGYG-PGQQGPGGYGPGOOGPSGPGSA 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAAGGAGQGGYG-GLGSQGTSGRGGLGGQGAGAAAAAAAAAAAG----GAGQGGYGGLGS 657
                                                                                                                                                                                                                                                             2 ASMTGGQQMGRGSMASGRGGLGGQGAG----AAAAAAAAAAAAGGGGGGGGGGGGGGTG-GLGSQGTSG 56
                                                                                                                                                                                                                                                                               PGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAG----GAGQGGYGGLGSQGTSGRGGLG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---GLGGQGAG-----AAAAAAAAAGGAGQGGYG------GLGSQGTS----
                                                                                                                                                                                                        Length 714;
                                                                                                                                                         ; OTHER INFORMATION: Description of Artificial Sequence:protein US-09-861-597-10
                                                                                                                                                                                                     38.0%; Score 1346.5; DB 10; Lengt
ilarity 50.0%; Pred. No. 1.3e-73;
Conservative 26; Mismatches 248; Indels
PRIOR APPLICATION NUMBER: 09/247,806
PRIOR FILING DATE: 1999-02-11
PRIOR APPLICATION NUMBER: EF 98/01614
PRIOR FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
                                                                                                                  TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OGTSGPGGYGPGQQ 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PGQQGPGGYGPGQQ 714
                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                     Query Match
Best Local Simi
Matches 367;
                                                                                                                                                FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                     171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             345
                                                                                                                                                                                                                                                                                                                                                                                                        129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            463
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                            ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

RESULT 6 US-09-861-597-2

```
GENERAL INCORANTE.

GENERAL INCORANTE.

GENERAL INCORANTE.

APPLICANT: PHILLIPPE, Michel

APPLICANT: ARADUBEAU, Jean-Claude

APPLICANT: ARADUBEAU, Jean-Pierre

TITLE OF INVENTION: COSMETIC ON DERMATOLOGICAL COMPOSITION CONTACTING AT

TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN

TITLE OF INVENTION: ANALOG

FILE REFERENCE: 6388-0385-0

CURRENT APPLICATION NUMBER: US/09/861,597

CURRENT APPLICATION NUMBER: US/09/861,806

PRIOR PELING DATE: 1999-02-11

PRIOR APPLICATION NUMBER: FR 98/01614

PRIOR APPLICATION NUMBER: FR 98/01614

SPRIOR PELLING DATE: 1998-02-11

NUMBER OF SEQ ID NOS: 14

SSOFTWARE: PATENTING VET. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GQGAGAAAAAAAAAAGGGGYGGLGSQGTSGRGGLG-GQ----GAGAAAAAAAAG 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    255 QGPGGYGP-GQQGPSGPG-----SAAAAAAAAAAAGPGGYG----PGQQGPGGYGP 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   401 AAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAAAAAGGGGYGGLGSQGTS 459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---PGQQGPGGYG-----PGQQGPGGYGPGSASAAAAAGPGQQGPGGYGP-GQQGPS 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GOQTSGRGGLG-----GQGAGAAAAAAAAAGGAGGGYGGLGSQGTSGRG----GLG 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142 GOG--GPGGYGPGQQGPSGPGSAAAAAAASGPGQQGPGGYGP-GQQGPGGYGPGQQGPS 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            341 GQQTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL--GGQGAGAAAAAAAAAAAGGAGQGGYG 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA 637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P-AQQGPSGPG------IAASAASAGPGGYG-----PAQQGPAGYGPGSAV 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  66 GAAAAAAAAAAGGAGQGGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAAAAGAGQGGYG-- 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           199 GPGSAAAAAAAAAAGPGQQGPGGYG----PGQQGPGGYGPGQQGLSGPGSAAAAAAAAAAGPGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               281 AGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGGGGYGGLGSQGTSGPGGYGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GPGQQGPGGYGPGQQGPSGPGSAAAAAAAA----AGPGGYG----PGQQGPGGYG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           299 GQQ------GPSGAGSAAAAAAAGPGQQGLGGYG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              31.0%; Score 1098.5; DB 10; Length 529; 44.4%; Pred. No. 4.9e-59; Live 25; Mismatches 159; Indels 203;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/09861597 Patent No. US20020064539A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; TYPE: PRT
; ORGANISM: Nephila clavipes
US-09-861-597-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 309; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

R APPLICATION NUMBER: 60/078939 R FILING DATE: 1998-03-20 R APPLICATION NUMBER: 60/079664 R APPLICATION NUMBER: 60/079786 R FILING DATE: 1998-03-27 R FILING DATE: 1998-03-27 R FILING DATE: 1998-04 15

R APPLICATION NUMBER: 60/082568

R FILING DATE: 1998-04-21

R APLICATION NUMBER: 60/082569

R FILING DATE: 1998-04-21

R APPLICATION NUMBER: 60/082704

R FILING DATE: 1998-04-22

R APPLICATION NUMBER: 60/082797

R APPLICATION NUMBER: 60/082797

R APPLICATION NUMBER: 60/08332 R APPLICATION NUMBER: 6/080194
R FILING DATE: 1998-03-31
R APPLICATION NUMBER: 60/080327
R FILING DATE: 1998-04-01
R RAPLICATION NUMBER: 60/080333
R FILING DATE: 1998-04-01
R APPLICATION NUMBER: 60/081049 R FILING DATE: 1998-04-08
R APPLICATION NUMBER: 60/081070
R FILING DATE: 1998-04-08
R FILING DATE: 1998-04-09
R FILING DATE: 1998-04-09
R APPLICATION NUMBER: 60/081838 R FILING DATE: 1998-04-29
R APPLICATION NUMBER: 60/083499
R FILING DATE: 1998-04-29
R FILING DATE: 1998-04-29
R R FILING DATE: 1998-04-29
R APPLICATION NUMBER: 60/084366
R FILING DATE: 1998-05-05 FILING DATE: 1998-04-28
APPLICATION NUMBER: 60/083495
FILING DATE: 1998-04-29
APPLICATION NUMBER: 60/083496 FILING DATE: 1998-05-07
APPLICATION NUMBER: 60/084640
FILING DATE: 1998-05-07
APPLICATION NUMBER: 60/084643
FILING DATE: 1998-05-07 APPLICATION NUMBER: 60/085700 FILING DATE: 1998-05-15 APPLICATION NUMBER: 60/086023 APPLICATION NUMBER: 60/086392 FILING DATE: 1998-05-06 APPLICATION NUMBER: 60/084639 APPLICATION NUMBER: 60/085573 FILING DATE: 1998-05-15 APPLICATION NUMBER: 60/085579 FILING DATE: 1998-05-15 APPLICATION NUMBER: 60/087098 PILLING DATE: 1998-05-28 APPLICATION NUMBER: 60/087208 FILLING DATE: 1998-05-28 APPLICATION NUMBER: 60/084414 APPLICATION NUMBER: 60/085580 60/085582 APPLICATION NUMBER: 60/086486 FILING DATE: 1998-05-28 APPLICATION NUMBER: 60/087609 1998-05-15 1998-03-20 FILING DATE: 1998-05-22 FILING DATE: 1998-03-3 FILING DATE: 1998-05-2 1998-05-1 APPLICATION NUMBER: FILING DATE: FILING DATE: PRIOR PRIOR

818

758

9

```
1297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1059 AAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATCAG 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1298 TTGAGGAAGCAACAGTCAAAGAAGGGAACTTCCCGG-GAAGATCAGACCCTTGCACTGCT 1356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1357 GAACCAGTTTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCC 1416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1417 TGAAACAGAAGTAGAAGATGATGAAG-GATGGATGTCACATGTACTTCAGTTTGAGGATA 1475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1476 AAAGCAGAAAAGTGAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATC 1535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1119 TCAGCCGCAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGG-GAACTCTTAGCA 1177
                                                                                                                    159
                                                                                                                                                                                                                                                                                                                                                                        879 AAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGTTGTAGAAA 938
                                                                                                                                                                                                                                                                                                                                                                                                                          939 GTGAAAAAGGTGATGCACCAGATTTAGTTGATGAGAAGATGAAAGTGCAGAGCATG 998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            273 AAAAAAGGAGGGYGGLGSQGTSGRG----GLGGQGAGAAA-----AAA 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   341 -----GQQTSGRGGLGGQGAGAAAAAAAAAGGA-----GQGGYGGLGSQGTSGRGGLG 389
                                                                                                                                                                                                                    160 YGGL----G$QGTSGPGG----YGPGQQTSGRGGLGGQGAG------AAAAAAA 199
                                                                                                                                                                                                                                                                                                                        AAAGGAGQGGYG------GLGSQGTSGRGGLGGQGAGAA 232
                  ----YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG 124
                                                                                                                       -----GGAGQGG
                                                                                                                                                    819 AAGCTGAGGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAGGCAAAAGCA
                                                                  699 CTTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAAAGAGAAACCAGAGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          590 -- LGG----- QGAGAAAAAAAAAAGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     390 GQGAGAAAAAAAAAAGGAGGGYGGLG~~~~~~SQGTSGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 443 GGAG-----QGGYGGLGSQGTSGRGG----LGGQGAGAAAAAAAAAAAGGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --GOGGYGG----LGSQGTSGPGGYGPGQQTSGRGGLGGQGA-----LGSQGTSGPGGYGPGQTSGRGGLGGQGA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----TSGRGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Sequence 2, Application US/09804898
; Patent No. US20020045264A1
; GENERAL INFORMATION:
; APPLICANT: IJURING, MATTHEW
; APPLICANT: XIAO, WEIDONG
; TITLE OF INVENTION: PRODUCTION OF CHIMERIC CAPSID VECTORS
                                                                                                                 ----LGSQGTSGRGGLGGQGAGAAAAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                525 -AAAAAAGGAG-QGGYGGLGSQGTSGRGGLGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAGAAA-----GAAGATAAAATGAGAATA 1619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         643 AAGGAGQGGYGGLGSQGTSGPGGYGPGQQTS 673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  563 AAAAAAGGAGQGGYGGLGSQG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
US-09-804-898-2
                                                                                                                                                                                                                                                                                                                        200 /
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1596
                                                                                                                 125
               84
                                                               9
                                                                                                                 à
                                                                                                                                                            Q
                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | | | | | ATGATGACGAGATAGATCCACACAAAATAAAAAGCTGTGAGGTTTTGTTTAATC 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 TGGQQMGRGSMASGRGGLGGQGAGAAAAAAAAAAAAAGAGGGGGGGGG--LGSQGTSG----- 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 393;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24.4%; Score 863.5; DB 12; Length 1894; 30.9%; Pred. No. 1e-44; 1ve 30; Mismatches 303; Indels 393;
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/08729
PRIOR APPLICATION NUMBER: 60/08025
PRIOR APPLICATION NUMBER: 60/08025
PRIOR APPLICATION NUMBER: 60/08028
PRIOR FILING DATE: 1998-06-04
PRIOR FILING DATE: 1998-06-04
PRIOR FILING DATE: 1998-06-04
PRIOR FILING DATE: 1998-06-04
PRIOR PRILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/08029
PRIOR PELING DATE: 1998-06-04
PRIOR PELING DATE: 1998-06-05
PRIOR PELING DATE: 1998-06-05
PRIOR PELING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/08020
PRIOR PELING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/08021
PRIOR PELING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/08021
PRIOR PELING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/08032
PRIOR PILING DATE: 1998-06-10
PRIOR PILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/08032
PRIOR PILING DATE: 1998-06-10
PRIOR PILING DATE: 1998-06-10
PRIOR PILING DATE: 1998-06-10
PRIOR PILING DATE: 1998-06-11
PRIOR PILING DATE: 1998-06-12
PRIOR PILING DATE: 1998-06-13

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 30.9%
Matches 325; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     633
```

g ò g

ò

```
565 AAAAGGAGGGYGGLGS-----QGTSGRGGLGGQG---AGAAAAAAA
                                                                                                                                                                                                                                                                                                                                      Sequence 269, Application US/10052586
Patent No. US20020127584A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILLING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION NUMBER: 60/066772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION NUMBER: 60/069335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: 1997-11-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR FILING DATE: 1997-12-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Goddard, Audrey
Godowski, Paul J.
Gurney, Austin L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pan, James
Smith, Victoria
                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
                                                                                                                                                                                                                          1047 -ATCCAGGAG 1055
                                                                                                                                                                                 658 QGTSGPGGYG 667
                                                                                           605
                                               g
                                                                                                                                 g
                                                                                                                                                                                 ò
                                                                                           ò
                                                                                                                                                                                                                                                                                                                                      36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          209 CGTGAGTGTTTTGCGACATTTTGCGACACCCATGTGGTCACGCTGGGTATTTAAGCCCGAG 268
                                                                                                                                                                                                                                                                                                                                                                                                                     94 AGCGAGCGAGAGCGCGCAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTGGAGGGGT 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100 -GQGAG-------AAAAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       269 TGAGCACGCAGGGTCTCCATTITGAAGCGGGAGGTTTGAACGCGCAGCCGCCATGCCGGG 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              329 GITITACGAGATTGTGATTAAGGTCCCCAGCGACCTTGACGAGCATCTGCCGGGATTTC 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       207 QGGYGGLGSQGTSGR---GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       389 TGACAGCTITGTGAACTGGGTGG-CCGAGAAGGAATGGGAGTTGCCGCCAGAITCTGACA 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              264 GQGA---GAAAAAAAAAAAGGAGQG------GYGGLGSQGTSGRGGLGGQGAGAAA 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       448 TGGATCTGAATCTGATTGAGCAGGCACCCTGACCGTGGCGGAGAAGCTGCAGCGGGACT 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  310 AAAAAAAGGAGGGYGGLGS---QGTSGPG---------GYGPGQQTSG 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           508 TICIGACGGAATGGCGCCGTGTGAGTAAGGCCCCGGAGGCCCTITICITGTGCAATTIG 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      347 RGGL-GGQGAGA------G 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             377 LGSQGTSGRGGLG-----GYG 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        628 TGGTTTTGGGACGTTTCCTGAGTCAGATTCGCGAAAACTGATTCAGAGAATTTACCGCG 687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             688 GGATCGAGCCGACTTTGCCAAACTGGTTCGCGGTCACAAAGACCAGAAATGGCGCCGGAG 747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      452 GLG-----SQGTSGRGGLGGQGAGAAAA-----------AAAA------479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      868 CGGAGCGTAAACGGTTGGTGGCGCAGCATCTGACGCACGTGTCGCAGACGCAGGAGCAGA 927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              141 AGAAAA-----PGGYGGYGGLGSQGTSG---PGGY 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  568 AGAAGGGAGAAGCTACTTCCACATGCACGTGCTCGTGGAACCACCGGGGTGAATCCA 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    414 GLGSQG---------TSGRGGLGGQGAGAAAAAAAAAAGGAG-QGGYG 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   523 ---aaaaaaaagagggggggggggggggggggagaaaa 564
                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     66
                                                                                                                                                                                                                                                                                                                                                                               2 ASMIGGQQMGRGSMASGRGGLGGQGAGAAAAAAAAAA------GGAGQGGY 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   47 GGLGSQGTSGRGGLGGQGAGAAAAAAAAAGG----AGQGGYGGLGSQGTSGRGGLG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            480 -----AGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLG---GQGAGAA----
                                                                                                                                                                                                                                                                                      ; Score 742; DB 10; Length 4679;
; Pred. No. 3.1e-37;
21; Mismatches 346; Indels 312;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  175 GPGQQTSGRGGLGGQGAGAAAAAAAAAGGAG-------
FILE REFERENCE: 102182-14
CURRENT PAPLICATION WOMBER: US/09/804,898
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/189,110
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin version 3.0
                                                                                                                                                                                                                      ; ORGANISM: adeno-associated virus US-09-804-898-2
                                                                                                                                                                                                                                                                                        20.9%;
30.0%;
                                                                                                                                                                                                                                                                                      Query Match 20.99
Best Local Similarity 30.09
Matches 291; Conservative
                                                                                                                                                      SEQ ID NO 2
LENGTH: 4679
                                                                                                                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    名
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
```

```
APPLICANY: Watanabe, COLIN K.
APPLICANY: Watanabe, COLIN K.
APPLICANY: Watanabe, COLIN K.
APPLICANY: Wacanabe, COLIN K.
APPLICANY: Wacanabe, COLIN K.
APPLICANY: Wacanabe, COLIN K.
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REPERENCE: P3430RL01
CURRENT APPLICATION NUMBER: 0202-01-15
PRIOR APPLICATION NUMBER: 060/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/063260
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-21
PRIOR FILING DATE: 1997-10-24
PRIOR FILING DATE: 1997-10-24
PRIOR FILING DATE: 1997-10-24
PRIOR PRILOR TON NUMBER: 60/063541
PRIOR PRILOR TON NUMBER: 60/063541
PRIOR PRILOR DATE: 1997-10-28
PRIOR FILING DATE: 1997-10-28
PRIOR PRILOR DATE: 1997-10-28
PRIOR PRILOR DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR APPLICATION NUMBER: 60/063541
PRIOR APPLICATION NUMBER: 60/063541
PRIOR PRILOR DATE: 1997-10-28
PRIOR PRILOR DATE: 1997-10-29
PRIOR PRILOR DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063541
PRIOR APPLICATION NUMBER: 60/063541
PRIOR PRILOR DATE: 1997-10-31
PRIOR PRILOR PAPELICATION NUMBER: 60/063541
PRIOR PRILOR DATE: 1997-10-31
PRIOR PRILOR PAPELICATION NUMBER: 60/063541
PRIOR PRILOR PAPELICATION NUMBER: 60/063541
PRIOR PRILOR DATE: 1997-10-31
PRIOR PRILOR PAPELICATION NUMBER: 60/06311
PRIOR PRILOR DATE: 1997-10-31
PRIOR PRILOR PAPELICATION NUMBER: 60/06311
PRIOR PRILOR DATE: 1997-10-31
PRIOR PRILOR PAPELICATION NUMBER: 60/06311
PRIOR PAPELICATION NUMBER: 60/06311
                                                                                                                                        928 ACAAAGAGAATCAGAATCCCAATTCTGATGCGCCGGTGATCAGATCAAAAACTTCAGCCA
                                                                                      ----AAGGAG-QGGYGGLGSQG-TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGS
```

```
APPLICATION TO THE STATE OF THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R APPLICATION NUMBER: 60/084414

R FILING DATE: 1998-05-06

R APPLICATION NUMBER: 60/084639

R FILING DATE: 1998-05-07

R APPLICATION NUMBER: 60/084640

R FILING DATE: 1998-05-07

R FILING DATE: 1998-05-07

R APPLICATION NUMBER: 60/086573

R RILING DATE: 1998-05-15

R APPLICATION NUMBER: 60/08559

R FILING DATE: 1998-05-15

R APPLICATION NUMBER: 60/08559

R FILING DATE: 1998-05-15

R APPLICATION NUMBER: 60/08559

R FILING DATE: 1998-05-15

R APPLICATION NUMBER: 60/085580

R APPLICATION NUMBER: 60/085580
NR FILING DATE: 1997-12-17

NR APPLICATION NUMBER: 60/068017

NR FILING DATE: 1997-12-18

NR APPLICATION NUMBER: 60/077450

NR FILING DATE: 1998-03-10

NR APPLICATION NUMBER: 60/077632

NR APPLICATION NUMBER: 60/077649

NR APPLICATION NUMBER: 60/077649

NR APPLICATION DATE: 1998-03-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 1998-04-22
APPLICATION NUMBER: 60/082797
FILING DATE: 1998-04-22
APPLICATION NUMBER: 60/083322
FILING DATE: 1998-04-28
APPLICATION NUMBER: 60/083495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R FILING DATE: 1998-04-29

R APPLICATION NUMBER: 60/083496

R APPLICATION NUMBER: 60/083499

R APPLICATION NUMBER: 60/083499

R APPLICATION NUMBER: 60/083559

R APPLICATION NUMBER: 60/083559

R APPLICATION NUMBER: 60/083559

R APPLICATION NUMBER: 60/083559

R APPLICATION NUMBER: 60/083559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LING DATE: 1998-03-27
PLICATION NUMBER: 60/079786
LING DATE: 1998-03-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FLICATION NUMBER: 60/082568
FPLICATION NUMBER: 60/082568
F-4s: DATE: 1998-04-21
                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: 60/078886
FILING DATE: 1998-03-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: 60/078939
FILING DATE: 1998-03-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CATION NUMBER: 60/081049
IG DATE: 1998-04-08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CATION NUMBER: 60/081070
IG DATE: 1998-04-08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 1998-05-15
APPLICATION NUMBER: 60/086023
FILING DATE: 1998-05-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PLICATION NUMBER: 60/079664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: 60/080107
FILING DATE: 1998-03-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: 60/080194
FILING DATE: 1998-03-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: 60/080333 FILING DATE: 1998-04-01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CATION NUMBER: 60/081195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CATION NUMBER: 60/081838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PLICATION NUMBER: 60/085700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CATION NUMBER: 60/080327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LING DATE:
                                                                                                                                  PRILOR PR
```

R APPLICATION NUMBER: 60/087098
R FILING DATE: 1998-05-28
R APPLICATION NUMBER: 60/087208
R FILING DATE: 1998-05-28
R FILING DATE: 1998-06-02
R FILING DATE: 1998-06-02
R APPLICATION NUMBER: 60/087759 APPLICATION NUMBER: 60/088740 FILING DATE: 1998-06-10 APPLICATION NUMBER: 60/088811 FILING DATE: 1998-06-10 FILING DATE: 1998-06-11 APPLICATION NUMBER: 60/089090 FILING DATE: 1998-06-12 FILING DATE: 1998-06-02 APPLICATION NUMBER: 60/087827 APPLICATION NUMBER: 60/088025 FILING DATE: 1998-06-04 APPLICATION NUMBER: 60/088033 FILING DATE: 1998-06-04 APPLICATION NUMBER: 60/088167 FILING DATE: 1998-06-10 APPLICATION NUMBER: 60/088738 FILING DATE: 1998-06-11 APPLICATION NUMBER: 60/088876 APPLICATION NUMBER: 60/089105 FILING DATE: 1998-06-12 APPLICATION NUMBER: 60/089512 FILING DATE: 1998-06-16 FILING DATE: 1998-06-17 APPLICATION NUMBER: 60/089653 FILING DATE: 1998-06-17 APPLICATION NUMBER: 60/086392 60/088028 APPLICATION NUMBER: 60/088029 APPLICATION NUMBER: 60/088202 APPLICATION NUMBER: 60/088212 60/088326 APPLICATION NUMBER: 60/088655 FILING DATE: 1996-00-09
APPLICATION NUMBER: 60/088722 APPLICATION NUMBER: 60/088824 FILING DATE: 1998-06-10 CATION NUMBER: 60/088825 APPLICATION NUMBER: 60/088826 APPLICATION NUMBER: 60/088861 FILING DATE: 1998-06-11 APPLICATION NUMBER: 60/088863 APPLICATION NUMBER: 60/089514 APPLICATION NUMBER: 60/089538 APPLICATION NUMBER: 60/089598 APPLICATION NUMBER: 60/089908 FILING DATE: 1998-06-04 FILING DATE: 1998-06-05 FILING DATE: 1998-06-05 1998-05-22 FILING DATE: 1998-06-04 1998-06-04 FILING DATE: 1998-06-03 FILING DATE: 1998-06-10 FILING DATE: 1998-06-FILING DATE: 1998-06-ICATION NUMBER: APPLICATION NUMBER: FILING DATE: 1

Query Match 19.4%; Score 687.5; DB 12; Length 1300; Best Local Similarity 28.3%; Pred. No. 2e-34; Matches 296; Conservative 17; Mismatches 341; Indels 391; Gaps

45;

28 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Ou Be Ma Qy Ob	Qy Oy Ob	64 64 64 64 64 64 64 64 64 64 64 64 64 6	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	60 60 60 60 60 60 60 60 60 60 60 60 60 6
				GAGGGGYGGLGSQGTSGRGGLGGQGAGA	
2 123 55 183 105 243 363 363	204 423 219 483	257 543 287 603	300 663 338 723 394 783	444 839 482 899 525 959	553 1018 604 1078 645
8 8 8 8 8 8 8 8 8	8 8 8 8	8 8 8	3 8 8 8 8	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	6 6 6 6 6

-09-060-8548-2
Sequence 2, Application US/09060854B
Sequence 2, Application US/09060854B
Sequence 2, Application US/09060854B
Sequence 2, Application US/09060854B
SEQUENCE 1NCORNATION:
APPLICANT: Estell, David Aaron
TITLE OF INVENTION: Human Protease and Use of Such Protease for Pharmaceutical
TITLE OF INVENTION: Applications and for Reducing the Allergenicity of No. US2002
FILE REFERENCE: GC532
CURRENT APPLICATION NUMBER: US/09/060,854B
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 2. 28; 99 GGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAG-- 156 185 GGCAGGGAAATCAAACGGGGAAA-AGAAATATTGTCGGGTTTAAACAGACAATGAGCA 243 157 QGGYG-GLGSQGTSGPGGYGPGQQTSG-----RGGLGGQGAGAAAAAAAAAGGA 205 264 GQGAGAAAAAAAAAGGAGQGGYGGLGGGGGGGGGGGGAGAAAAAAAGGGGG 323 358 AATTGAAAAAAGACCCGAGCGTCGCTTACGTTGAAGAAG---ATCACGTAGCACATGCGT 414 509 SGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA--GAAAAAAA 566 531 CATCCTGATTTAAAGGTAGCAAGCGGAGCCAGCATGGTTCCTTCTGAAACAAATCCTTTC 590 431 -- AGAAAAAA---- AAAGGAG------QGGYGGLGSQGTSGRGGLGGQGAGAAAAAA 477 125 GCTGTTTGCTTTAGCGTTAATCTTTACGATGGCGTTCGGCAGCACCATCCTCTGCCCAGGC 184 puery Match
19.1%; Score 675.5; DB 10; Length 1497;
sest Local Similarity 30.7%; Pred. No. 1.1e-33;
fatches 265; Conservative 22; Mismatches 330; Indels 245; Gaps 244 CGATGAGCGCCCCTAAGAAGAAAGATGTCATTTCTGAAAAAGGCGGGAAAGTGCAAAAGC 471 CAAGGCTACACTGGATCAAATGTTAAAGTAGCGGTTATCGACAGCGGTATCGATTCTTCT 80 ------RGGLGSQ--GTSG------RGGL 304 AATTCAAATATGTAGACGCAGCTTCAGTCACTTAAACGAA----AAAGCTGTA--AAAG 409 -----QGGYGGLGSQGTSG-RGGLGGQG-----478 AAAGGAG----QGGYG-----YGPGQQT 651 ATCGGTGTATTAGGCGTTGCGCCAAGCGCATCACTTTACGCTGTAAAAGTTCTCGGTGCT ORGANISM: B. amyloliquefaciens 09-060-8548-2 Query Match Best Local Similarity LENGTH: 1497 TYPE: PRT

```
GENERAL INFORMATION:
APPLICANT: GUEGLER, Karl et al.
TITLE OF INVENTION: SOLOGIATED HUMAN PHOSPHOLIPASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REPERBENCE: LOLOGO845DIV
CURRENT APPLICATION NUMBER: 60/232,632
PRIOR PILING DATE: 2002-03-14
PRIOR FILING DATE: 2000-09-14
PRIOR FILING DATE: 2000-09-14
PRIOR FILING DATE: 2000-09-14
PRIOR FILING DATE: 2000-12-18
                                                                                                                                                                                         1473 CCCTGCAGGCCTCGCCACTGCAGATCCTCGTGAAGGAGGAGGGCCCCCGGGCCGGGTCCT 1532
                                                                                                                                                                                                                                     1353 GCTCCACGCCCCCCGTGTCTCCCACCCCCTCGGAGCGCTCACTGCTCAGCACGGGCGATG 1412
                                                                                               1413 AAAACTCCACCCCCGGGGACACCTTTGGTGAGATGGTGACATCACCTCTGACGCAGCTGA 1472
                                                        -AAAAAAAAGGAGQGGYGGLGSQGTSG----RGÇLGGQGAGAA 472
                                                                                                                                                                                                                                                                                                                                    576 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGA--GQGGYGGLG-----SQGTSGRGG 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            431 ACCAGGACAGCAAGATGAGCTTCAAGGAGATCAAGAGCCTGCTGAGAATGGTCAACGTGG 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----GAGQG------GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAG-QGG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 551 GTCTAGAGGGGCTGAGATCGAGGAGTTCCTGCGGGCGGCTGCTGAAGCGGCCGGAGCTGG 610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122 YGGLGS-----GYGGLGGQGAGAAAAAAAAAAAGGAGQG-----GYGGLGSQG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           611 AGGAGATCTTCCATCAGTACTCG-GGCGAGGACCGCGTGCTGAGTGCCCCTGAGCTGCTG 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GAGTTCCTGGAGGACCAGGGGGGGGGCGCCACACTGGCCCGGGCCCCAGCACCTCATT 729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GAGQGGYGGLGSQGTSG-RGGLGGQGAAAAAAAAAGGAGQ 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         191 ACATGAACGACATGTACGCCTACCTCCTCTTCAAGGAGTGTGACCACTCCCAACAACGACC 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 ASMTGGQQMGRGSMASGRG----GLGGQGAGAAAA-----AAAAAGG----AGQGG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           46 YGGLGSQ-GTSGRG-----AAAAAAG--- 77
                                                                                                                                           473 AAAAAAAAGAGOG-----GYGG-----LGSQGTSGPGGYGPGQQTSGRGGLGGQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                              628 LGGQGAGAAAAAAAAAGGAGQGG------YGGLGSQGTSGPGGYGPG 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         168 TSG----PGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAAG-----PGGYGPGQTSGRGGLGGQGAGAAAAAAAAAAAAAAAAAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 18.7%; Score 662; DB 9; Length 2211; Best Local Similarity 27.3%; Pred. No. 9.6e-33; Matches 281; Conservative 22; Mismatches 355; Indels 372;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/10096961
Patent No. US20020155572A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; ORGANISM: Homo sapien
US-10-096-961-1
                                                   431 AGAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
US-10-096-961-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    204
                                                                                                                                                                                                                                                                                                                                                                                qq
                                                   ö
                                                                                            g
                                                                                                                                               ò
                                                                                                                                                                                         g
                                                                                                                                                                                                                                       á
                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 33;
                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: OLSON, ERIC N.
APPLICANT: OLSON, ERIC N.
APPLICANT: OLSON, ERIC N.
APPLICANT: WANG, DA-ZHI
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO A CARDIAC-SPECIFIC
TITLE OF INVENTION: NUCLEAR REGULATORY FACTOR
FILE REFERENCE: OTSD:6950S
CURRENT APPLICATION NUMBER: US/10/029,217A
CURRENT APPLICATION NUMBER: 08/10/029,217A
PRIOR APPLICATION NUMBER: 60/257,761
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                945 CCACTACCAATAGCAGCTCCAGCTCGGCCCCTGGGCCCTGTGG-GCTGGCACGTCAG 1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1004 AACAGCACCTCACTGACTGGCAAGCCG--GGAGCCCTGCCGGCCAACCTGGACGACGATGA 1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CCAAGGCCCCTGCCGCCACTCTATCCTGCACAAGGCTGGCGAGGTGGTG--GTAGCCTT 1239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1240 CCCAGCGGCCCGGCTGACGCGGGCCCAGCCCTGGTGGCAGGCCTGGCTCCAGCTGA 1299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1062 AGGTGGCAGAGCTGAAGCAGGAGCTGAAGTTGCGATCACTGCCTGTCTCGGGCACCAAAA 1121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1122 CTGAGCTGATTGAGCGCCTTCGAGCCTATCAAGACCAAATCAGCCCTGTGCCAGGAGCCC 1181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            831 GTTGATAAAGCCGTTGCATCCGGCGTCGTAGTCGTTGCGGCAGCCGGTAACGAAGGCACT 890
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --AAAAAAAAAAGGAG------QGGY-----GGLGSQGTSGRGGLGGQGAGAAA 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         157 QGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGA-------GAAAAAAAAA 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   202 AGGAGQGGY---GGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGG---YGGLGSQG 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --RGGLGG-----QGAGAAAAAAAGGAGQGGYGGLGS-----QGTSGP 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----AAAAAAAAAAGGAGQG---GYGGLGSQGTSGRGGLGGQG 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26 GAGAAAAAAAAAAGAGGG---YGGLGSQGTSGRGG-----LGGQGAG----- 66
    --LGSQGTSGRGGLGGQG
                                           771 ATGGACGTTATTAACATGAGCCTCGGCGGACCTTCTGGTTCTGCTGTTTAAAAGCGGCA
                                                                                          595 AGAAAAAAAAAAGGAG----QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             256 TSGRGG----LGGQGAGA-----AAAAAAAAAAGGAGQGGYGGLGSQGTSG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19.0%; Score 672; DB 9; Length 3907; llarity 30.9%; Pred. No. 3.7e-33; Conservative 15; Mismatches 342; Indels 22(
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
    ----GOGGXGG---
                                                                                                                                                                                                                                                                                                                                                   Sequence 24, Application US/10029217A
Patent No. US20020164735A1
GENERAL INFORMATION:
                                                                                                                                                                                      -----AGQGGYGGLGSQGTSG 662
                                                                                                                                                                                                                                  891 TCCGGCAGCTCAAGCACAGTGG 912
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                               US-10-029-217A-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; ORGANISM: Mus
US-10-029-217A-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Sim
Matches 258;
567 AAGGA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1300
                                                                                                                                                                                         647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 395
                                                                                                                                   a
                                                                                                                                                                                                                                q
à
                                           셤
                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

37;

Gaps

```
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KAY, Brian K.
FOWLKES, Dana M.
MCCONNELL, Stephen J.
INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1086 CGACCATGCCTTCACGCTGTCCCTTACCCTGTCATCCTATCCCTGGAGAACCACTGCGG 1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1146 GCTGGAGCAGCA--GGCTGCCATGGCCCGCCACCTCTGCACCATCCTGGGGGACATGCTG 1203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1204 GTGACACAGGGGCTGGACTCCCCAAATCCCGAGGAGCTGCCATCCCCAGAGCAGCTGAAG 1263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1317 CCGGGCTCTGTCGGATCGGGAGGAGGGGGGGGGAGGATGACGAGGAGGAGAAGAGGAGGT 1376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1377 GGAG-CCTGCAGCGCCGCTGGCCAAGCAGATCTCCCCGGAGCTGTCGCCCTGG 1435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1026 CTATCATGGCCATACCCTCACCTCCAAGATTCTCTTCCGGGACGTGGCCCAAGCCGTGCG 1085
                                                                                                                                  845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----GLGSQGTSG---RGGLGGQGAG 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGYGGLGSQGTSGRGGLGGQGA------GAAAAAAAAAAAAAGGAG---QGGYG 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      582 -----QGTSG----RGGLGGQGAGAAAAAAAAAGGAGQGGYGG-LGSQGTSGRGGL 628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         664
                                                                                                                                                                                                                                                              846 GTTCCAGGACATGAACCAGCCCCTTGCCCACTACTTCATCTTCCTCCCACACAACACCTA 905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     THE STATE OF THE S
730 CAGACCTATGAGCTCAACGAGACAGCCAAGCAGCATGAGCTGATGACACTGGATGGCTTC 789
                                                                                                                                                                                                                                                                                                                                316 ------AGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAA 361
                                                                 247 GYGGLGS---QGTSGR----GGLGGQGA------GAAAAAAAAAAGGAGQGGYGGLGS
                                                                                                                           790 ATGATGTACCTGTTGTCGCCGGAGGGGACTGCCTTGGACAACACCCACACGT----GTGT
                                                                                                                                                                                                                                                                                                                                                                                906 TCTGACTGACTCCCAGATCGGGGGCCCAGCAGCAGCCAGAGGCCTTTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    490 GLGSQGTSGPGGYGPGQQT-----SGRGGLGGQG-----AGAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA...---AAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGQGAGAAAAAAAAAAGGAGQGG-----PG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   362 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HOFFMAN, NO. US20020034755A1h
                                                                                                                                                                                                292 QGTSGRGGLGGQGAGAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 37, Application US/09879957 Patent No. US20020034755A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SPARKS, Andrew B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: New York
STATE: New Yor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1436 CTGTGTACTG 1445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     665 GYGPGQQTSG 674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-879-957-37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      532 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     629
윱
                                                                                                                           셤
                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                 ð
                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  76 AGGAGQGGYGGLGSQGTSGRGGLGG------QGAGAAAAAAAAAAAAGGAGQG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               109 -----GCTGTCCAGGCACCCTGGTCCACTGCAAAAAGGGTCCACTTACCATTTCTGCA 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               148 ----AAAAAGGAGQGGYGGLGSQGTSGPGG------YGPGQQTSGRGG---L 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 284 CCCG-GCTGGCTTGGAGGAATTAAAAGGAAAGACAGG----GTGGTTCCCTGCAAACT 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----QGAGAAAAAAAAAAGGAGGGYGGIGSQGTSGRGGIGG---QGAGAAAAAAA 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51 GGAAGCACAAGACAAGCTGGGTCGGCTTTTCCATCAACACCAAGAACCAGCTAAGCCA-- 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    164 CAGGAAATGTAAAAGTGGTGTATTACCGGGCACTGTACCCCTTTGAATCCAGAAGCCAT 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            224 GATGAAATCACTATCCAGCCAGGAGACATAGTCATGGTGGATGAAGCCAAACTGGAGAA 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        339 ATGCAGAGAAAATCCCCAGAAAATGAGGTTCCCGCTCCAGTGAAACCAGTGACTGATTCAA 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 344; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16 ASGRGCLGGQGAGAAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                247 GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    359 AA-------AAAAAAGGAGQGGYG-GLGS-------QGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     159 CTTCAGAGCCCTCCACGACCCCTAATAACTGGGCCGACTTCAGCTCCACGTGGCCCACCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            519 GCACGAATGAGAAACCAGAAACGGATAACTGGGATGCATGGGCAGCCCAGCCCTCTCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                299 GLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   399 CATCTGCCCCTGCCCCCAAACTGGCCTTGCGTGAGACCCCCGCCCCTTTGGCAGTAACCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               383 ----SGRG------GLGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 1400;
                               COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: S/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <UNKNOWN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 GYGGLGSQGTSGRGG----LGGQ-----GAGAAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
18.4%; Score 652; DB 10;
Best Local Similarity 28.3%; Pred. No. 2.7e-32;
Matches 274; Conservative 24; Mismatches 325;
                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US 08/630,915
FILING DATE: 03-APR-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                    NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE DESCRIPTION: SEQ ID NO: 37: US-09-879-957-37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SQ IN 37:

SEQUENCE CHARACTERISTICS:

LENGTH: 1400 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELECOMMUNICATION INFORMATION:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRANDEDNESS: single
                                                                                                                                                                                                                                               PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    391
```

```
APPLICANT:
APPLICANT:
APPLICANT:
                   Query Match
                                                                                                                                                                                                                                                                173
                                                                                                                                                                                                                                                                                                                                                                                                                                                 330
                                                                                                                                                                                                                                                                                                                           225
                                                                                                                                                                                                                                                                                                                                                         208
                                                                                                                                                                                                                                                                                                                                                                                      280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   629
                                                                                                       g
                                                                                                                                                                     a
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                          ò
                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
   -AGAAAAAAAA-- 479
                                                                                      638 GGCTCCTCCCGTCTCCTGTAGGCCAG--GGTGAAAGGTGGAGGGGGTACAAGCTC 695
                                                                                                                                                    696 AAGCCCTATATCCTTGGAGAGCCAAAAAGACAACCACTTAAATTTTAACAAAAATGATG 755
                                                                                                                                                                                  ------GGAGQGGYGGLGSQGTSGRGGLGGQGAGA---AAAAAAAAGGAGQ 573
                                                                                                                                                                                                    GGYGGLG----SQGT-----SGRGTGGGGGAGAAAAAAAAAGGAG 610
                                                                                                                                                                                                                                                                             GTTGGTTCCCCAAGTCTTACGTGAAACTCATTTCAGGGCCCCATAAGGAAGTCTACAAGCA 875
                                                                                                                                                                                                                                                                                                                                      TGGATTCTGGTTCTTCAGAGAGTCCTGCTAGTCTAAAGCGAGTAGCCTCTCCAGCAGCCA 935
                                                                                                                                                                                                                                                                                                          611 QGG----YGGLGSQGTSGRGG---LGGQGAGAAAAAAAAAAAAGGAG-----QGGYGGLGS 657
                              579 CCGTTCCAAGTGCCGGC-CAGTTAAGGCAGAGGTCCGCCTTTACTCCAGCCACGCCACT
                                                              ----AGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION UMBER: US/09/756,071B
FILING DATE: 08-Jan-2001
CLASSIFICATION: <UNKNOWN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kallunki, Pekka
Pyke, Charles
TITLE OF INVENTION: Laminin Chains: Diagnostic Uses
NUMBER OF SEQUENCES: 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS:
ADDRESSEE: Fay Sharpe Fagan Minnich & STREST: 1100 Superior Ave, Suite 700 CITY: Cleveland STATE: Ohio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATORNEY/AGENT INFORMATION:
NAME: MIDITCH, RICHARD, J.
REGISTRATION NUMBER: 24,175
REFERENCE/DOCKET NUMBER: TRV 20014
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US 09/663,147
                                                                                                                         -----LGGQGAGAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 20:
15-09-756-071B-20
441 AAG----GAGQGGYGGLGSQGTSGRGGLGGQG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 20, Application US/09756071B Patent No. US20020052307A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEPHONE: 216-861-5582
TELEFAX: 216-241-1666
INFORMATION FOR SEQ ID NO: 20:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 720 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Tryggvason, Karl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: USA
ZIP: 44114
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                 AGCCGAG 942
                                                                                                                                                                                                                                                                                                                                                                      658 QGTSGPG 664
                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
US-09-756-071B-20
                                                                                                                                                                                                                                               574
                                                            480
                                                                                                                         514
                                                                                                                                                                                                                                                                           816
                                                                                                                                                                                                                                                                                                                                                                                                 936
                                                                                                                                                                                  531
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                g
ò
                              셤
                                                                                      g
                                                                                                                     ò
                                                                                                                                                    셤
                                                                                                                                                                                  ò
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                    ð
                                                            ð
```

```
35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     426
                                                                                                                                                                                                                                                               LGSQGTSGRGCLGGQGAGAAAAAAAAAGAGGGGYGGLGS-----QGTSGPG---- 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    487 TGCCTCTGGAGGCTGGGTCCTCCTTATTCACAGGTGAGTCACACCCTGAAACACAGGCT 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               574 --GGYGGLGSQGTSGRG-GLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 630
                                                                                                                                                                                                                                                                                                                                                                                                114 AGAGGTT-----GACCCTAAACTGGAG-GGTGGAGGGACCCTGTTGTGACTCTC 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                           224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GAGQGGYGGLGSQGT----SGRGGLGGQG---AGAAAAAAAAAAAGGAGQGGYGGLGS 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                389 GGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGGA 448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GYGGLG-SQGTSGRGGLGGQGAGAAAAAAAAAG---GAGQGGYG----GLGSQGTSG-- 498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              427 GGGGAGCGGCGGCGGCGGGGAGCGATTTTCCAGCCCGGTTTGTGCTCTGTGTTTTGTC 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       547 CTCTTCCTGTCAG-GACTGAGTCAGGTAGAAGAGTCGATAAAACCACCTGATCAAGGAAA 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       606 AGGAAGCCACAGCGGAGCGCAGAGTGAGAACCACCAACCGAGGCGCCC-----GGCAG 658
                                                          Gaps
                                                                                                                                                                                                                            76 AGGAGQGGYGGLGSQGTSGRGG---LGGQGAGAA-----AAAAAAAGGAGQGGYGG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          163 CGACTGACTTGTCTTCCTTGATG----TCCTTTAAGCCGGA------GCTGATTC 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGGCTGCTGCTTATTTCTGAGTTAGCGCTCTTAAGATTGGGCCTCCCAGTTTGAGGAAG 267
                                                                                                               16 ASGRGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 75
                                                                                                                                                                    8 ATAGGGAGGTTGGCCAGTCAATAGGT-----TACTTTATGAGTTG----CTAACCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----AGQGGYGGLG--SQGTSGRGGLGGQGAGAAAA-----AAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        -GYGPGQQTSG------RGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               376 TCTGTGCCACACCCTTGGCCCG -- GGCCAGGT ----- GTGCGCCCTCCTCGCTGCGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -------GCTCCGGGGAATCTCGCACATTCCAGGCAAAGGCTCCCGGGCCGCAGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----PGGYGPGQQTSGR------GGLGGQGAGAA-----AAAAAAGG--
                                                          187;
  Length 720;
18.4%; Score 651.5; DB 10; Length 32.5%; Pred. No. 1.7e-32; Live 17; Mismatches 314; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CGACCCCTGCAGCGGAGACAGAGACGGGCCCGGCACCGCCATG 705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               631 QGA------GAAAAAAAAAAAGGGGGG--YGGLGSQGTSGPG 664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Sequence 525, Application US/10052586; Patent No. US20020127584A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pan,Jāmes
Smith,Victoria
Watanabe,Colin K.
Wood,William I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Desnoyers, Luc
Goddard, Audrey
Godowski, Paul J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gurney, Austin L.
                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Baker, Kevin P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chen, Jian
                             Best Local Similarity
Matches 249; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-10-052-586-525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT:
```

```
R FILING DATE: 1998-04-08
R APPLICATION NUMBER: 60/081195
FILING DATE: 1998-04-09
R PILING DATE: 1998-04-15
R FILING DATE: 1998-04-15
R FILING DATE: 1998-04-1
R R FILING DATE: 1998-04-21
R APPLICATION NUMBER: 60/082569
R FILING DATE: 1998-04-21
R APPLICATION NUMBER: 60/082569
R FILING DATE: 1998-04-21
R APPLICATION NUMBER: 60/082569
R FILING DATE: 1998-04-21
                                                                                                                                                                                  R FILING DATE: 1998-04-28

R FILING DATE: 1998-04-28

R FILING DATE: 1998-04-28

R RILING DATE: 1998-04-29

R RILING DATE: 1998-04-29

R RILING DATE: 1998-04-29

R RILING DATE: 1998-04-29

R RPLICATION NUMBER: 60/083495

R FILING DATE: 1998-04-29

R RILING DATE: 1998-05-06

R RILING DATE: 1998-05-06

R RILING DATE: 1998-05-06

R RILING DATE: 1998-05-07

R RELING DATE: 1998-05-15

R RELING DATE: 1998-05-16

R RELING DATE: 1998-06-02

R RILING DATE: 1998-06-02

R RILING DATE: 1998-06-03

R RILING DATE: 1998-06-03

R RILING DATE: 1998-06-03

R RILING DATE: 1998-06-04

R RELING DATE: 1998-06-04

R RELING DATE: 1998-06-04

R RELING DATE: 1998-06-05

R RILING DATE: 1998-06-05

R RILING DATE: 1998-06-07

R RELING DATE: 1
          R APPLICATION NUMBER: 60/06331

R FILING DATE: 1997-11-13

R APPLICATION NUMBER: 60/066420

R FILING DATE: 1997-11-21

R APPLICATION NUMBER: 60/066422

R FILING DATE: 1997-11-24

R APPLICATION NUMBER: 60/06672

R FILING DATE: 1997-11-24

R APPLICATION NUMBER: 60/069425

R FILING DATE: 1997-12-11

R APPLICATION NUMBER: 60/069425

R FILING DATE: 1997-12-12

R APPLICATION NUMBER: 60/069425

R FILING DATE: 1997-12-13

R APPLICATION NUMBER: 60/069870

R FILING DATE: 1997-12-18

R APPLICATION NUMBER: 60/07450

R FILING DATE: 1998-03-10

R APPLICATION NUMBER: 60/077632

R FILING DATE: 1998-03-11

R APPLICATION NUMBER: 60/078939

R FILING DATE: 1998-03-11

R APPLICATION NUMBER: 60/078939

R FILING DATE: 1998-03-12

R APPLICATION NUMBER: 60/078939

R FILING DATE: 1998-03-20

R APPLICATION NUMBER: 60/078939

R FILING DATE: 1998-03-27

R APPLICATION NUMBER: 60/080327

R APPLICATION NUMBER: 60/080337

R FILING DATE: 1998-03-31

R FILING DATE: 1998-04-08

R FILING DATE: 1998-04-08

R FILING DATE: 1998-04-08

R FILING DATE: 1998-04-08
```

```
G--LGGQGAGAAAAAAAAAGGAGQGG-----YGGLGSQGT----SGRGGLGGQGA- 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----GAAAAA-----AAAAAGGAGQGGYG-------GLGSQGTSGRGGL 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      508 TGGAGGAGTGCCTGCTCAGACCCAGGCTGACACAGGAACTCTCCCAGAAG----GCTG 563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------GLGGQGAGAAAAAAAAGGAGQGG 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           564 CAGGCTTCCCTCCCAACCCTCCAACGCCCACTTGCTGAGCCCGTGCTTATCTGTTCAG 623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              272 GAATGGAGAATTCCAACCCAGGAGGAGACTGTGCCTGGGGCTGCCACGAGAAGTGGTGAC 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   332 ACTICGAGAAACGGGTICCAAATGCACGTIGCCTICCAGTICTCTGTGTGACCTIGGGCA 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     392 AGTCACTTCAGCTCCCTGAGCTTTGTTTTAAAATATTTTTAAAATGTATAAAACCATG 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GACCATTACATATGAAGAGAAATGT----GTGTGCAAACATTCAGTTAATAATCACAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GG------ABABABGCAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAG------GAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     YGGLGSQGTSGRGGLGGQG----AGAAAAAAAAAGG-----AGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 4440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17.2%; Score 610; DB 12;
llarity 26.3%; Pred. No. 1.9e-29;
Conservative 28; Mismatches 342;
IR APPLICATION NUMBER: 60/088217
IR FILING DATE: 1998-66-05
IR APPLICATION NUMBER: 60/088326
IR APPLICATION NUMBER: 60/08855
IR APPLICATION NUMBER: 60/088728
IR FILING DATE: 1998-06-09
IR FILING DATE: 1998-06-10
IR FILING DATE: 1998-06-10
IR APPLICATION NUMBER: 60/088740
IR FILING DATE: 1998-06-10
IR APPLICATION NUMBER: 60/08811
IR FILING DATE: 1998-06-10
IR APPLICATION NUMBER: 60/08824
IR FILING DATE: 1998-06-10
IR APPLICATION NUMBER: 60/08826
IR FILING DATE: 1998-06-10
IR APPLICATION NUMBER: 60/088861
IR FILING DATE: 1998-06-10
IR APPLICATION NUMBER: 60/08861
IR FILING DATE: 1998-06-11
IR APPLICATION NUMBER: 60/08976
IR FILING DATE: 1998-06-12
IR APPLICATION NUMBER: 60/08976
IR FILING DATE: 1998-06-12
IR APPLICATION NUMBER: 60/089512
IR APPLICATION NUMBER: 60/089518
IR FILING DATE: 1998-06-16
IR APPLICATION NUMBER: 60/089518
IR FILING DATE: 1998-06-16
IR APPLICATION NUMBER: 60/089518
IR FILING DATE: 1998-06-17
IR APPLICATION NUMBER: 60/08958
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR PLING DATE: 1998-06-04
PRIOR PLING DATE: 1998-06-04
PRIOR PLING DATE: 1998-06-04
PRIOR PLING DATE: 1998-06-07
PRIOR APPLICATION NUMBER: 60/08872
PRIOR PLING DATE: 1998-06-10
PRIOR PAPLICATION NUMBER: 60/08874
PRIOR FILING DATE: 1998-06-10
PRIOR PLING DATE: 1998-06-10
PRIOR PAPLICATION NUMBER: 60/08861
PRIOR PLING DATE: 1998-06-11
PRIOR PAPLICATION NUMBER: 60/08876
PRIOR PLING DATE: 1998-06-11
PRIOR PAPLICATION NUMBER: 60/08976
PRIOR PLING DATE: 1998-06-12
PRIOR PAPLICATION NUMBER: 60/08976
PRIOR PLING DATE: 1998-06-12
PRIOR PAPLICATION NUMBER: 60/08951
PRIOR PLING DATE: 1998-06-12
PRIOR PAPLICATION NUMBER: 60/08951
PRIOR PLING DATE: 1998-06-16
PRIOR PAPLICATION NUMBER: 60/08951
PRIOR PLING DATE: 1998-06-17
PRIOR PAPLICATION NUMBER: 60/08958
PRIOR PLING DATE: 1998-06-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       256 TSGRG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local ST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

GGTAAAGGTTGCTGAATAATCAGGCTGCTGGTTAGTTTTTACATTTCACCTTTCCCAGTG 1043 1153 TTACTAATTATGAATTCCAGTGCTTTCCTTGCTGAAAGAGAGGCGTGGAATCAACGCTGA 1212 AAGGACAGAAAGGTCTTCTGTTTGCACCTACCCTAAGGATTTGGGGTAGACACTGGGAAT 1152 570 647 804 GAAATGGAGGTCACGCAGTGGAAATAGGATCCAGACAGATTAATCCAATCTCAAGCCTGA 863 334 367 GAGQGG-YGGLGSQGT-----AAAAA 404 GTAAGGTCAGTACTGTCACTTCCCCCACTTTGAAGATGAGGAGAGCACAAATTCTAGATG 803 864 ATTCTTCCATTCCACGCTACGCTTGAAGCTCAATCTCTTCCTGGTTGATTCTCCCCAC 923 924 IICCCCACCCCAGAIAIAICCCAICGCIGCIIGGIGGACAGIAGCCAIGACIGGGIIII 983 --GLGGQGAGAAAAAAAAAGGAGGGGGGGGGGGGGGGG PGGYGPGQQTSGRG-----ALG-GQGAGA-------AAAAAAAAAG 624 TGGAAATGCCATGTGCCACGTTCCTCAGAAGTAAAGCCACTAGGTAAGTGTGACACTCTG ------LGSQGTS-----GRGCLGCQGAGAAAAAAAAAAGGAGQGGY RGGLGGQGAĞAAAAAAAAAAAAAAAAAAAAAAAAAGG 1044 AAATGGGGCCCCATGAAAAGGCA----GCTCAAGTTGT-----AAATTACTCAAAGG 571 AGGGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAGGAGQ -----GGYG-----GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGA GGLGSQG------TSGPGGYGPGQYCPCQT-----SG-1213 GIGAAGGCAICAAGIIIAAG 1232 GOGGYGGLG -- SQGTSGPGG 665 286 YGGLGSQGTSGRG--744 511 368 443 489 984 1093 453 g g g ð g à 셤 g g 셤 ò g ò ò ð ç õ à

Search completed: December 18, 2002, 16:39:20 Job time: 28.5457 secs

37;

507 255 us-09-490-291-6.rpr

Page 1

```
GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2002, 16:13:49 ; Search time 20.4171 Seconds

(without alignments)
3253.588 Million cell updates/sec
```

Title: US-09-490-291-6
Perfect score: 3543
Sequence: 1 MASMTGGQQMGRGSMASGRG.....TSGIRIRAPSTSFEHHHHH 691
Scoring table: BLOSUM62

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0 Maximum DB seq length: 2000000000 Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	major ampullate f1		tic	hypothetical glyci	[ca]	spidroin 2, dragli	ica]	hypothetical glyci		hypothetical glyci			_		_	precu	elastin precursor		hypothetical glyci										
SUMMARIES	ID	A36068	T31328	F70806	D70807	B70807	A70934	E70806	A70869	H70846	E70917	A70896	A44112	A70812	H70987	B70812	F70963	E70895	E70835	D70835	E70824	C70974	A70807	G70917	F70824	F70620	EAMS	EART	E70820	A70893
	DB	N	~	~	N	~	~	7	~	7	~	~	~	7	7	~	N	~	~	7	7	~	7	~						
	Length	718	2639	1901	1489	1079	1306	1381	1660	1538	1329	853	627	749	914	882	778	167	837	957	783	731	714	741	801	101	860	864	923	299
d	Query	60.3	39.1	36.2	35.6	35.1	33.8	33.8	33.3	32.7	32.5	32.5	32.1	31.5	31.3	30.6	29.6	29.3	29.1	29.1	28.6	28.4	28.0	27.9	27.9	27.6	27.5	26.5	26.4	26.1
	Score	2138	1386	1281.5	1260.5	1242	1198.5	1196.5	1180	1158.5	1152	1151	1138.5	1114.5	1110.5	1083	1049.5	1038.5	1032.5	1032	1014	1001	992	686	988.5	977.5	916	939	937	923.5
	Result No.		7	m	4	S	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29

අ **දු** අ

hypothetical glyci hypothetical glyci hypothetical glyci	glycine-rich cell glycine-rich cell hypothetical prote hypothetical glyci		hypothetical glyci hypothetical glyci hypothetical glyci	elastin precursor, hypothetical glyci hypothetical glyci	חמריפתו מחורואפו פם
F70868 G70846 A70770	S01820 T31611 A70900	G70804 S59623 A70989	D70931 H70816 B70523	EABO F70971 H70589 OOBE31	+ 63970
000	400	000	000	- 777	4
694 484 603	465 1585 576	584 770 618	639 606 591	747 588 615	5
พหม	باوي	-, 00, 00,	~ 6 6	מטטת	,
252	23.34	223	222	22222	4
903.5 893 869	853 848.5	817.5 809 808.5	804 801 800	786.5 777.5 776.5)
30 31 32		36 37 38	39 41 41	4 4 4 4 0 6 6 6	,

ALIGNMENTS

la clavibes) (fragme text_change 09-Sep. dragiine silk. 2494 9712 sidue 292 as Gin, G	08 2; Length 718; 1-113; 49; Indels 184; Gaps	SQGAG-AAAAAAAAAGGAGQ 	YGGLGSQGTSGRGGLGGGGA 	RGGLGGQGAGAAAAAAA RGGQGAAAAA	AGGAGGGGYGGLG 	-AAAAGGAGGGYGGLGS 	GGYGP-GQQTSGRGGL : GGYGGLGSQGAGRGGL	LGSQGTSGRGGLGGQG
RESULT 1 A36068 major ampullate fibroin protein - orb spider (Nephila clavipes) (fragment) C;Species: Nephila clavipes C;Species: Nephila clavipes C;Saccession: A36068 R;Xu. M.: Leavis, R.V. Brixu. M.: Leavis, R.V. A;Reference number: A36068; MulD:90384959; PMID:2402494 A;Reference number: A36068; MulD:90384959; PMID:2402494 A;Status: preliminary A;Accession: A36068 A;Status: preliminary A;Accession: A36068 A;Status: preliminary A;Reference number: A3707137; NID:9159711; PID:9159712	Query Match 60.3%; Score 2138; DB 2; Best Local Similarity 66.3%; Pred. No. 3.6e-113; Matches 498; Conservative 20; Mismatches 49;	25 QGGGAQAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGGG	82 GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAGGGGGGGGGGGGGGGGGGGGG	142 GAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAA 201	202 AGGAGGGGYGGLGSOGTSGRGGLGGGGAAAAAAAAAAAGGAGGGGYGGLG 111111111111111111111111111111111111	253 SQGTSGRGGLGGQCAGAAAAA	292 QGTSGRGGLGGQGAGAAAAAAAAGAGGGGGGGGGGGGGGGGGGG	351 GGGGAGAAAA
ULT 1 De am De cie ate: At	sery sat L	0	യഹ	14	20	25	29	35
RESULT A36068 major a C; Spec C; Datec C; Acce R; Xu, b Proc. I A; Resle A; Resle A; Resle A; Resle A; Resle A; Resle A; Rocies A; Rocie	ōmx	oy B	oy g	Qy Dp	Qy Db	Oy Op	Qy Pp	o o

##############################
FEBSULT 2 1913128

A; Experimental source: strain H37Rv C; Genetics: A; Gene: Rv3514 C; Superfamily: collagen alpha 1(IV) chain

```
H37RV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1054 ATGAAGLGDNGGVGGDGGAGGAAGNGGNAGVGLTAKAGDGGAAGNGGNGGAGGAGGAGDN 1113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               994 GGGNGGRGGDGGDGASGLGLGLSGFDGGQGGQGGAGGSAGAGGINGAGGAGGNGGDGGDG 1053
                                                                                                                                                                                                                                                                                                               355
                                                                                                                                                                                                                                                                                                                                                                       393
                                                                                                                                                                                                                                                                                                                                                                                                   816
                                                                                                                                                                                                                                                                                                                                                                                                                                435
                                                                                                                                                                                                                                                                                                                                                                                                                                                          876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SINANGGAGGNGGTGGKGGAGGAGTLGVGGSG---GTGGDGGDAGSGGGGGFGGAAGKAG 993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     580
                                                      459
                                                                                    168
                                                                                                                                           207
                                                                                                                                                                     578
                                                                                                                                                                                                                              636
                                                                                                                                                                                                                                                        311
                                                                                                                                                                                                                                                                                                                                           756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------AAAAAAAGGAGQGGYGGLGSQG----TSGRGGLGGQGA----GAAAAA 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           581 SQGTSG---RGGLGGQ-GAGAAA------AAAAAAAGGAGQGGYGGLGSQG-- 621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DATATGATGGTGFSGGAGGAGGAGGNTGVGTNGSGGQG-GTGGAGGAGGAGGAGGADNPT
                                                                                                                                                                                                                340 GDGVLGGVGGTGGKGGVGGVAGLGGAGGAAGQLFSAGGAAGAVGVGGTGGQGGAGGAGAA
                            G------TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGR--
                                                                                                            460 ASGIGADGGAGGTGGNAGAGGAGGAGGAGGTGGTGGVVGAAGKAGIGGTGGQGGAGGAGSAGT
                                                                                                                                           -----SGPGGYGPGQTSGRGGLGGQGAGAAAAAAAAAAAAA
                                                                                                                                                                                                 --GGYGGLGSQGTSGRGGLGGQG--AGAAAAAAAAGG-AGQGGYGGLGSQGT---SGR
                                                                                                                                                                                                                                                       GGLGGQGAGAAAAAAAAAGGAGQGGYGG-----LGSQGTSGRGGLGGQ-GAGAAAA
                                                                                                                                                                                                                                                                                                              AAAAAGGA----GQGGYGGLGSQ----GTSGPGGYGPGQQTSGRGGLGGQGA-----
                                                                                                                                                                                                                                                                                                                                                                       ----GAAAAAAAAAAGGA-GQGGYGGL-----GSQGTSGRGGLGGQGA-----
                                                                                                                                                                                                                                                                                                                                                                                        ------GAAAAAAAAAAAAAAAAAAAAA---RGGLGGQ-GAGAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                         817 FDGGOGGOGGAGGSAGAGGINGAGGAGGOGGDGATGAAGLGDNGGVGGDGGAGGAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                877 NGGNAGVGLTAKAGDGGAAGNGGNGGAGGAGGAGDNNFNGGQGGAGGQGGQGGLGGASTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAAAAGGA----GQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  532 GAGQGGYGGLGSQGTS-----GRGGLGGQGAGAAAAAAAAAAAGGGGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --TSGRGGLGGQGA----GAAAAAAAAAAGGA----GQGGYGGLGSQGTSGPGGYG 667
                                                                                 -----GGTGGQG--AGAAAAAAAAGG------
                                                                                  134
                                                                                                                                                                    520
                                                                                                                                                                                                                                                       260
                                                                                                                                                                                                                                                                                                              312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   622
                                                                                                                                         169
                                                                                                                                                                                                208
                                                                                                                                                                                                                                                                                                                                                                       356
                                                                                                                                                                                                                                                                                                                                                                                                                              394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            476
                                                                                                                                                               g
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           윱
유
                            õ
                                                g
                                                                                 õ
                                                                                                            셤
                                                                                                                                        ò
                                                                                                                                                                                                ò
                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                               οχ
                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
hypothetical glycine-rich protein Rv3514 - Mycobacterium tuberculosis (strain H37RV)
c)Species: Mycobacterium tuberculosis
c)Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
c)Accession: D70807
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davles, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A,Authors: Squares, R.; Sullston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A,Authors: Squares, R.; Sullston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A,Reference number: A70500; MUID:98295987; PMID:9634230
A,Reference number: A70500; MUID:98295987; PMID:9634230
A,Status: preliminary; nucleic acid sequence not shown; translation not shown
A,Residues: 1-1489 <Col.>
```

hypothetical glycine-rich protein Rv3512 - Mycobacterium tuberculosis (strain H37RV)

RESULT B70807

A;Cross-references: GB:AL022022; GB:AL123456; NID:g3261554; PIDN:CAA17751.1; PID:g29244\$

```
37;
                                                                                                                                                                                            403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                871 CGQVVAGGAGISFSNGSNGGTGGTGGVG--CTGGDGGNAGTGAGDDGKGGTGGTG-GTGG 927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------GTSGRGGLGG----QGAGAAAAAAAAAGGAGQGGYGGLGSQGTS 131
                                                                                                                                                                                                                                                                                               227
                                                                                                                                                                                                                                                                                                                                                                                                                256 -TSGRGGLGGQGAGAAAAAAAAAAGGGGGGGGGG-LGSQGTSGRGGLGGQ----GAGAAA 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        691
                                                                                    214 GGGAGGVGGAGGGTGGAGGRAELLFGAGGAGGAGTDGGPGATGGTGGHGGVGGDGGWLAP 273
                                                                                                                                                                                                                                                                 392 GDGGAGGAGADADQPGATGGTGFAGGAGGAGGSSGAGGTNGSGGAGGGGGGAGGAGA 451
                                                                                                                                                                                                                                    132 GRGGLGGQGA------GAAAAAAAAAAAAGGA------GQGGYGGLGSQGTSGPGGYG
                                                                                                                                                                                                                                                                                                               274 GGAGGAGGQGGAGGAGGALGGTGGTGGTGGAGGAGGRGALLLGAGGQGGLGGAGGQG
                                                                                                                                                                                                                                                                                               176 PGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG-LGSQGTSGRGGLGGQ-----
                                                                                                                                                                                                                                                                                                                                                                                   512 GAGADADQPGATGGTGFAGGAGGAGGAGGSSGAGGTNGSGGAGGTGGQGGAGGAGGAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AGGAGQGGYGGLGSQGT---SGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGG-LGSQGTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GVGGAGGOGGDGGAGADADQPGATGGTGFAGGAGGAGGSGGSSCAGGTNGSGGAGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --------GQGGYGGLGSQGTSGPGG-YGPGQQTSGRGCLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          524 AAAAAAAGGAG-QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    583 GTSGRGGLGGQGAGAAAAAAAAAGGA-----GQGGYGGLGSQGTSGRGGLGGQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                986 SAGGTGGTGGAGGNAGAGGLANTGGTAGNAGIGGDGGQGGNGGQGDSG-SGLGGQPGFAG
                                                                                                                   36 AAAGGA-GQGGYGGLGSQG--TSGRGGLGGQ-GAGAAAAAAAAAGGAGQGGYGGLGSQ-
                                                                                                                                                                                                                                                                                                                                                         GAGAAA------AAAAAAGGA------GQGGYGGLGSQG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA-----AAAAGGAGQ-------GGYGGLGSQ-GTSGPGGYGPGQQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----GAGAAAAA-----AAAAAGGAGQGGYGGLG-----SQGTSGPGGYG 667
   Length 1489;
                             Indels
                                                          6 GGQQMGRGSMASGRGGLGGQ----GAGAAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      460 GRGGLGGQ----GAGAAAAA------AAAAGGA------
   DB 2;
                .1e-64
35.6%; Score 1260.5;
38.5%; Pred. No. 7.1e
Live 35; Mismatches
                             Conservative
               Similarity
               Best Local Sim
Matches 344;
Query Match
                                                                                                                                                                             93
                                                                                                                                                                                                                                                                                                                            452
                                                                                                                                                                                                                                                                                                                                                         228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             632
                                                          ð
                                                                                     셤
                                                                                                                   ò
                                                                                                                                             ద
                                                                                                                                                                             ç
                                                                                                                                                                                                    g
                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
RGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    115 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       265
                   994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            408
                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                              Ricole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A; Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A; Reference number: A70500; MUID:98295987; PMID:9634230
A; Reference number: A70500; MUID:98295987; PMID:9634230
A; Residues: DNA
A; Residues: 1-1079 <COL>
A; Residues: GB:AL022022; GB:AL123456; NID:93261554; PIDN:CAA17749.1; PID:9292444
A; Experimental source: strain H37Rv
C; Genetics: Rv3512
C; Superfamily: collagen alpha 1(I) chain; fibrillar collagen carbovul-toward.
                                                                                                                                                                                                                                                                                                                                                29;
Species: Mycobacterium tuberculosis
Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                               GAAAAAAAAAAGG-AGQGGYGGLGSQ--GTSGRGGLGGQGAGAAAAAAAAAGGAGQGGY 450
                                                                                                                                                                                                                                                                                                                                                                                                                                         54 TSGRGGLGGQG-----AGAAAAAAAAAAAAAGGA-GQGGYGGLGSQGTSGRGGLGGQ- 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       102 -----GAGAAAAAAAAAAAAAAAAAAAAA 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            176 GKGGDGGAGVTSSTAGNSGGAGGSGGKGGDAGAGGAGTPGANGIAGNGGDGGD--GAAG 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GSGNGGNGGSGGNGGNAGMGGNSGTGSGDGGAGGNGGAAGTGGTGGTGGLTGTGGTGGTGGSG 765
                                                                                                                                                                                                                                                                                                                                                                           5 TGGQQMGRGSMASGRGGLGGQGAGAAAAAAAAAAAGGA-GQGGYGG------LGSQG 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                234 AVGISGATGAGDGGHGGTGAAGGNGGTGGAGGSGIDGVGGGTGGTGGNGGNGAIGGAGGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG-----A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       245 QGGYGGLGSQGTSGRGGLGGQG-----AGAAAAAAAAAAAAAAGGAG------QGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQTSGRGGLGGQGAGAAA-AAAAAAAGGAGQGGYG-GLGSQGTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GANGGAGGAGGAGGTGGNGGAGGDAGNAGNGGNGTGNGGNGGNGGTAGMGGNGGAGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AGGAG-QGGYGGLGSQ------GTSGRGGLGGQG--AGAAAAAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGLGSQGTSGRGLGGQGAGAAAAAAAAAAGGGGGGGGGGGG-----SQGTSGPGGYGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAAAAGGAGGGYGGL----GSQGT-SGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSG 674
                                                                                                                                                                                                                                                                                                               35.1%; Score 1242; DB 2; Length 1079;
40.7%; Pred. No. 6.1e-63;
Live 34; Mismatches 306; Indels 120;
                                                                                                                                                                                                                                                                                                                              Best Local Similarity 40.78
Matches 316; Conservative
C;Species: Mycobacter
C;Date: 17-Jul-1998 (
C;Accession: B70807
R;Cole, S.T.; Brosch,
                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
Ayobst

Ayobst

Ayobst

Ayobst

C;Species: Mycobacterium tuberculosis

C;Accession: A70934

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

C;Acomor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamilin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

A,Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whithened, S.; Barrell, B.G.

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Status: preliminary: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1308 «COL.

A;Residues: 1-1308 «COL.

A;Residues: 1-1308 «COL.

A;Genetian

A;Genetian

A;Genetianily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAAAAAAAAAGG-AGQGGYGG--LGSQGTSGRGGLGGQGA------GAAAAAAAAA 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       147 --AAAAAAGGAGGAGGIGSQGTSG--PGGYG-PGQ--QTSGRGGLGGQGAGAAAAAA 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                637 GASGGRGGDGGLVGNGGAGGAGGNGSKGLPGLGRLGNPGLDGGTGGNGGAGGSGGAWAGN 696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      934
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --SGPGGYGPGQQTS 509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  365 AAGGA--------GQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAA 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 33.8%; Score 1198.5; DB 2; Length 1306; Best Local Similarity 38.4%; Pred. No. 1.9e-60; Matches 323; Conservative 37; Mismatches 293; Indels 189; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            200 AAAGG------AGQGGYGGLGSQ------GTSGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       235 AAAAAAGG-AGQGGYGGLGSQ------GTSGRGGLGGQG--AGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGTGGAGGTGGVGGTGGSGSDGVNGSSAGADGHPGGTGGVGGTGGKGGDGGDGGAAPNGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           815 GGAGPAGTTGSQGAGGNGGSGGTGGDPGDGGNGANGSVFTNNGIGGNGGNGGNAGPSGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               875 GSGGAGSTFGATGSSSSIHVNGGNGGNGGNGDHALSGNGAAGGNGGNGGNGSLRGSGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GQGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAGGAGQGGYGG-----LGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11 GRGSMASGRGGLGGQGAGAAAAAAAAAAGG----AGQGGYGGLGSQGTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GG------AGQGGYGGL------GSQGTSGRGGLGGQGA------GAAAA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             278 AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAG-GAGQGGYGGLGSQ-GTSGP
--GTGGDGGNGGNGADNTANMTAQAGGDGGNGGDGGFGGGAGAGGGGLTAGANGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGYGP-----GQQTSGRGGLGGQG-----
```

= = = =

us-09-490-291-6.rpr

Db 995 DGGVGGGGSVAGDGADGGRGGAGGTGGTGLRGTTGATGATGTFDAGADGHGGNGGTG 1052	
QY \$10 GRGGLGGQGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	DD 547 PGGAGGAGGAG
559 AAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGG	Qy 400 AAAAAGGAGGGGY
Db 1111 TLFSSLAGTGGTGGTGTGTGTGGAGGTGSTLGATGATGAAGRAGNGGVGGSG 1168	UD SOAN INAFVIGEEGGAGGDIGGA
Qy 619 SQGTS-GRGGLGGQGAGAAAAAAAAGGAGGGGGGGGSQGTSGPGGY-GPGQQT 672	DD 649 KAGTGNSGNFGVDGEAG
Qy 673 SG 674	
Db 1228 AG 1229	709
RESULT 7	QY 545GTSGRGGL Db 769 GKAPHGDAGGSGGDGGK
6/08/0 hypothetical glycine-rich protein Rv3507 - Mycobacterium tuberculosis (strain H37RV) C:Species: Mycobacterium tuberculosis	Qy 585 SGRGGLGGQGAGA
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000 C:Accession: E70806	Db 829 TITVPGNGGNAGDGGN-
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.	Oy 641 AAAAGGAG-QGGYGGLG
Nature 393, 537-544, 1998 A:Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome	RESULT 8
A; Meterence number: A/USUU; MUID: 9629396/; PMID: 903423U A; Accession: E70806	A/0869 hypothetical qlycine-rich p
A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA	C; Species: Mycobacterium tu C; Date: 17-Jul-1998 #sequen
A; Residues: 1-1381 <col/> A; Residues: 1-1381 <col/> A; Cross-references: GB: ALO22022; GB: AL123456; NID: 93261554; PIDN: CAA17744.1; PID: 929244	C; Accession: A70869 R; Cole, S.T.; Brosch, R.; P
A;Experimental source: strain H37Rv C;Genetics:	; Connor, R.; Davies, R.; D Rajandream, M.A.; Rogers, J
A;Gene: Rv3507 C;Superfamily: collagen alpha 1(IV) chain	Nature 393, 537-544, 1998 A;Authors: Sqares, R.; Suls
Query Match 33.8%; Score 1196.5; DB 2; Length 1381;	A; Title: Deciphering the bi A; Reference number: A70500;
Best Local Similarity 38.7%; Pred. No. 2.5e-60; Matches 319; Conservative 36; Mismatches 276; Indels 193; Gaps 36;	A; Accession: A70869 A; Status: preliminary; nucl
QY 7 GQQMGRGSMASGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTS 55	A; Molecule type: DNA A; Residues: 1-1660 <col/>
	A; Experimental source: stra
56 GRGGLGCOGAGABABABABABAGGGCGCYGGLGSQGTSGRGGLGCQGAGA	C; venetics: A; Gene: Rv2490c C; Superfamily: collagen alp
ZIU GNGGNGALLIGGGGLGGAGGMGGTGGGTGGTGGNGALLIGAGGVGGAGGTGGQGTGA	Query Match
UY 100 AAAAAAAAGGACGGGGGGGGGGGGGGGGGGAAAAAAAAA	Best Local Similarity 3 Matches 326; Conservati
162 GLGSOGTSGPGGYGPGOOTSGPGGLGGOGGGAAAAAAAAAAAAAAGGAGGGGGGGGGGGG	Qy 2 ASMTGGQQMGRGSMASG
326 GTGGAGGAGPVLFGHGGAGGMGGOGGTGGMGGAGGDGTTVI	Db 514 SSGTPGEDGGNGG-AGG
219 - SGRGGLGGGGAAAAAAAAAGGGGGGGGGGGGGGGGGGGGG	Qy 46 YGGLGSQCTSGRGGLGG
Db 367 AAGTGGEGGTGGAAGAAGAAGAAGALTSGGLAGGVGAGGTGGTGGTGGTGGTGGTGAAVGFG 426	
QY 277 AAGGAGGGGGGGGGG 1	Db 630 AFKDGDGGAGGDGGDPG
304GAGAAAAAAAAAAGGAGOGGYGGIGSOGTSGDGGYG	Qy 148 AAAAAGGAGQ
487 GGIGGAGGAAGTGNGGHACNTGDGGDGGTGGNGANGTGGVNGANNTLNPDTPGGAG	Db 690 ANGGKGGAGGNGGLVGD
	_

```
Laton, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
siology of Mycobacterium tuberculosis from the complete geno
j; MUID:98295987; PMID:9634230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         021246; GB:AL123456; NID:g3261507; PIDN:CAA16067.1; PID:g279 rain H37Rv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein Rv2490c - Mycobacterium tuberculosis (strain H37RV)
tuberculosis
ance_revision 17-Jul-1998 #text_change 20-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  leic acid sequence not shown; translation not shown
                                                                                                                                                                                                                                                                                                                                    828
                       SYGGL-----GSQGTSGRGGLGGQGAGAAAAAAAAAAGGGGG 449
                                                                                                 --- RGGLGGQGAGAAAAAAAAAAGGA----GQGGYGGLGSQGTSG 498
                                                                                                                                                                          3GLGGQGAGAAAAAAAAGG----AGQGGYGGLGSQ----- 544
                                                                                                                                                                                                                                                584
                                                                                                                                                                                                                                                                                                                                                                           GAAAAAAAAAAGGGGYGGLGSQGTSGRGGLGGQGAGAAAA 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGGAGAAAAAAAAAGGAGGGYGGL-GSQGTSGRGGLGGQGAG 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GQGGYGGLGSQG-----TSGRGGLGGQGAGAAAA 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QGGYGGLGSQGTSGPGGYGPGQQ------TSGRG 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   33.3%; Score 1180; DB 2; Length 1660;
37.6%; Pred. No. 2.4e-59;
11ve 38; Mismatches 292; Indels 212; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SGRGGLGGQGAGAAAAAAAAAGGAG-------0GG 45
                                                                                                                                                                                                                                                                                                 GSQGTSGPGGYGPGQQTSGIRIRAPSTS 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           pha 1(IV) chain
```

Oy 185 GLGGGGAGAAAAAAAGGAGGGGGGGGGGGGGGGGGAGAAAAA	
Qy 245 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGA-GQGGYGGLGSQGTSGRGGLGGQ 303	OY 59 GLGGGGA-GAAAAAAAGGGGGGGGGGGGGGGG
304 GAGAAAAAAAAGQGGYGGL-GSQGTSGPGGYGPGQQTSGRGG	Qy 102 GAGAAAAAAAAGGAGGGGGGGGGGGGGGGGGGGGGGG
LGG-QGAGAAAAAAAAAGGAGGGGGGGGGGGGGG	Oy 142GAAAAAAAAAAAGGGGGGGGGCGGTGGGGGGGGGG
388 LGGOGAGAAAAAAAAGGAGOGGYGGLGSOG-TSGRGGLGGO -GAGAAAAA 1	OY 189 GGAGAAAAAAAGGSAGOGGYGGLGSOGTSGRGGLGGOGAGAAAAAAAAAAGGGG 246
Qy 438 AAAAAGGACQCGYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAAGGAGGGGY 488	QY 247 GYGGLGSQGTSGRGGLGGQ-GAGAAAAAAAAGGGGGGGGG
Qy 489 GG	QY 289 1GSQGTSGRGGLGGQ
	QY 304GAGAAAAAAAAGGAGOGGYGGLGSOGTSGPGGY 338
TSGRGGLGGQGAGAAAAAAAAAGGA-GQGGYGGLGSQGTSGRGGLGGQGA 1	QY 339 -GPGQQTSGRGGLGGQCAGAAAAAAAAGGGCGCGCGCGCGCGCGGGGGGAGAA 397
GAAAAAAAAAAAGGAGQGYGGLGSQGTSGRGGLGGQGAGAAA-AAAAAAGGAGGGGGGGGGGGGGGGGGG	QY 398 AAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGGGGAAAAAAAAAGG 444
647	OY 445AGOGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGGGYGGLGSO 494
RESULT 9 H70846	QY 495 GTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLG 542
hypothetical glycine-rich protein Rv3345c - Mycobacterium tuberculosis (strain H37RV) C:Species: Mycobacterium tuberculosis C:Species: Mycobacterium tuberculosis Schee: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000 C:Accession: H70846	OY 543 SQGTSGRGGLGGQGAGAAAAAAAAAGGAGGGGGGGGGGGGGGGG
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 303, 877-544	OY 599 AAAA 637 I
A;Authors: Squres, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A;Reference number: A70500; MUID:98295987; PMID:9634230 A;Accession: H70846	Oy 638AAAAAAAGGAGGGGYGGLGSOGTSGPGGYGPGQQTSG 674 1
A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA A;Molecule type	RESULT 10 E70917 hypothetical diveine-rich protein Rv1450c - Mycobacterium tuberculosis (etrain ul7bv)
	C; Species: Mycobacterium tuberculosis C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000 C;Accession: E70917 R;Cole: S T · Brosch, R : Parkhill, I · Carnier T · Churcher C · Harris D · Cordon
ouery Best Match	Accordor, R.; Davies, R.; Deviln, K.; Feltwell, T.; Gentles, S.; Hamili, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998 A Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A;Tille: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
Qy 11 GRGSMASGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRG 58	A;Reference number: A70500; MUID:98295987; PMID:9634230

H37RV)

```
Species: Mycobacterium tiberculisis

C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 24-Nov-1999

C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 24-Nov-1999

C; Accession: A70896

R; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Rajandream, A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Rajandream, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A. Attle: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A. Reference number: A70500; MUID:98295987; PMID:9634230

A. Status: preliminary; nucleic acid sequence not shown; translation not shown

A. Mesidues: 1-853 < COL.>
A. Cross-references: GB-AL021897; GB:AL123456; NID:g3256022; PIDN:CAA17207.1; PID:e125
                                                                                                                                                        hypothetical glycine-rich protein Rv1091 - Mycobacterium tuberculosis (strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GAAAAAAAAAAAAAGGGGYGGLG-----SQGTSGRGGLGGQGAGAAAAAAAAAAGGA 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GQGCYGGLG--SQGTSGPGGYGPGQQTS----GRGGLGGQGA----GAAAAAAAAAGGA 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              311 GSGGTGGAGGSTAGAGGNGGAGGGGGTGGLLFGNGGAGGHGAAAGNGLAAGNGVSSSGGG 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTGGLLFGAGGNGGPGGSGGAADIG--GNGGAGNGGGTDGNGGNGGSGGGG---AGSGGDG 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GODGGPGGLLYGNGGNGGTSTTAGVAGGNGGAAGLIGNGGAGGGGGGAGAAGGNGGAGGWL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GQGGYGGLGSQGTSGRGGLGGQG-----AGAAAAAAAAAAAGGAGQGGYGGL-----GSQ 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        371 GAGGTGGAGGDG - GAGGAGGNARLWGVGGAGGAGGDGGAGGAGGAGGKGGSGLSGNANGGAG 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    540 NGGNGGNGGNGGNGATGGWLYGNGGAGGOGATAGAGGAGANGVSSTNGGGTGGNGGLG-- 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GTSGRGGLGGQGA--GAAAAAAAAAGGAG-QGGYGGLG----SQGTSGRGLGGQ-GAG 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7 GQQMGRGSMASGRGGLGGQGAGAAAA----AAAAAAGGAGQGGYGGLGSQGTS----- 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -GRGGLGGQGAGAAAAAAAAAAGGA-----GQGGYGGLGSQGTSG----RGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                : Score 1151; DB 2; Length 853;
; Pred. No. 6.3e-58;
27; Mismatches 307; Indels 110;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Gene: Rv1091
C;Superfamily: unassigned collagens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                32.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 40.49
Matches 301; Conservative
                                                          1022 ---PAGSILH 1028
                      677 IRAPSTSFEH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  658
                                                        g
                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
              A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA A;Residues: 1-1329 <COL>
A;Residues: 1-1329 <COL>
A;Cross-references: GB:Z95844; GB:AL123456; NID:g3250713; PIDN:CAB09271.1; PID:g2131046 A;Experimental source: strain H37Rv C;Genetics: A;Gene: Rv1450c
                                                                                                                                                      C; Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology
                                                                                                                                                                                                                                    35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :|| || || || | :
SQNGGAGGAGGAGGAGGDFGGAGGAGGAGGANGANTTPGAKGGDGGHGGPGAQGGN 729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            903
                                                                                                                                                                                                                                                                                                                                                                                      193 GNGGVGGAGGOSLLG-GATGGAGGNAGLFGVGGTGGPGGPGGPGGVG--GTGGAGGLGGT 249
                                                                                                                                                                                                                                                                                                                                                                                                                            GAGAAAAAAAAAAAAAGG-AGQGGYGGLGSQ----GTSGRGGLGGQGAGAAAAAAAAAAAA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                  --AAAAGGAGQGGYGGLGS--QGTSGPGGYGPGQQTSGRGGLG------GQGAGAAAA 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              310 TSGGVGGDAGDGGTGGRGGMLAGAGGAGGAGGVGGTGGAGGAGFSRALIVAGDNGGDPGA 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                430 GAGGAGGNGAPGAPPSGGDPNGGGGGAGGAGGKGGDGGAQAGDGGAGGKGGNGGNGA 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        490 TGATGLNGLGAGADGTDGGKGGNGCAGGGGGGGGGGKALAATHQDGSMGAGGAGGNGGA 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGLGGQG-----AGAAAAAAAAAAGGGGGGGGG-LGSQGTSGPGGYGPGQQTSGIR 676
                                                                                                                                                                                                                                                                                                          GOAGGAGGILWGNGGAGGSGAPGQVGGAGGAGLFCTGGAGGAGAGAAGGAGSGGWLL 192
                                                                                                                                                                                                                                                                                                                                                904 AGGNGGLGGFGGNGAAGANGVAVGAPGQPGGAGGHGGAGGNGGAGGNGGQGVVSDGAGGA 963
                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                      GYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAGGA-GQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --GLGSQGTSGPGGYGPGQQ-----TSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TSGRGGLGGQGAGAAAAAAAAAAGGAGQ----GGYGGLG------SQGTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                197 AAAAAAGGAGQ--GGYGGLGSQGTS-GRGGLGGQGAGAAAAAAAAAAAAGGA------GQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    370 GGAGGTGGAGSTIGAHGAAGASPTSGGNGGAGGNGAHFSSGGKAGGNGGAGGAGGLVGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -GQGGYGGLGSQGT-----SGPGGYGPGQQTSGRGGLGGQGA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AGQ----GGYGGLGSQ-----GTSGRGCLGGQGAGAAAAAAAAAAAAAAAGG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          730 GGQGGPGGLAGNLFGQNGIQGVGSGGKGGAGGLAGDGGNGANGNFAFGDGNGGHGGNGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              790 NPGAGGGGGGGGGGTPGAKGAHGFTPTSGGDGGDGGNGGN----SQVVGGNGGDGGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----AAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          845 GNGGSAGTGGNGGRGGDGAFGGMSANATNPGENGPNGNPGGNGGAGGAGGAGLNG-GNGG
                                                                                                                                                                                        Query Match 32.5%; Score 1152; DB 2; Length 1329; Best Local Similarity 35.8%; Pred. No. 7.6e-58; Matches 326; Conservative 35; Mismatches 305; Indels 244;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GL-GSQCTSGRGGLGGQGAGAAAAAA----
A; Accession: E70917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                610
                                                                                                                                                                                                                                                                                                          133
                                                                                                                                                                                                                                                                                                                                                42
                                                                                                                                                                                                                                                                                                                                                                                                                          102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          490
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                        å
                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                        οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              & 8
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

28;

QY 531 GGAGGCCYGGLCSOGTSGRGCLGGQCAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGCL 590 111 1 1 1 1 1 1 1 1 1 1	370 GP-GQQGPSGPGSASAAAAAAAGPGGYGP-GQQGPGCYAPGQQGPSGPG 471 AAAAAAAAAAGGAGQGYGGLGSQGTSGPGGYGPGGYGFGGGGGGGAAAAAAAA : :
RESULT 12 A44112 spidroin 2, dragline silk fibroin - orb spider (Nephila clavipes) (fragment) N:Alternate names: silk fibroin, dragline C;Species: Nephila clavipes C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Mar-1998 C;Accession: A44112; S27824 R;Himman, M.B.; Lewis, R.V. J. Blol. Chem. 267, 19320-19324, 1992 A;Title: Isolation of a clone encoding a second dragline silk fibroin. Nephila clavipes A;Reference number: A44112; MUID:92406876; PMID:1527052 A;Accession: A44112	DD 501 AGYGPGSAVAASAGAGSAGYGP-GSQASAAASRLASPDSGARVASAVSNLVSSGPT 555 Qy 648 GQGGYGGLGSQGTSGPGGYGPGQQTSGIRIRA 679 L
A;Status: preliminary A;Molecule type: mRNA A;Readdues: 1-627 cHLN> A;Cross-references: GB:M92913; NID:g159713; PID:g159714 A;Cross-references: GB:M92913; NID:g159713; PID:g159714 A;Note: sequence extracted from NCBI backbone (NCBIP:113893) A;Humman, M.B.; Lewis, R.V. submitted to the EMBL Data Library, May 1992 A;Bescription: Isolation of a clone encoding a second dragline silk fibroin: Nephila cla A;Reference number: S27824 A;Accession: S27824 A;Residues: 19-627 cHI2> A;Residues: 19-627 cHI2> A;Residues: 19-627 cHI2> A;Cross-references: EMBL:M92913	Ricole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamiln, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature. 393, 537-544, 1998 A; Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A; Aitle: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A; Reference number: A70500; MUD:98295987; PMID:9634230 A; Reference number: A70500; MUD:98295987; PMID:9634230 A; Status: preliminary; nucleic acid sequence not shown; translation not shown A; Residues: 1-749 <col/> A; Residues: 1-749 <col/> A; Residues: 1-749 <col/> A; Resperimental source: Strain H37Ry
Query Match 32.1%; Score 1138.5; DB 2; Length 627; Best Local Similarity 47.0%; Pred. No. 2.5e-57; Matches 325; Conservative 36; Mismatches 190; Indels 141; Gaps 36; Qy 23 GGGGGAAAAAAAAAGGAGGGGGGGGGGGGGGGGGAAAAAA	C;Genetics: A;Gene: Rv0833 C;Superfamily: elastin C;superfamily: elastin Outery Match Best Local Similarity 39.6%; Pred: No. 6.36-56; Matches 303; Conservative 18; Mismatches 309; Indels 135; Gaps 32;
79 AGGGYGGLG-SQGTSGRGGLGGGGGAAAAAAAAAGGGGGYGGLGSQGTSGR	Qy 18 GRGGLGGQGAGAAAAAAAAAAAAAAGGGGGGGGGGGGGGG
188 -GOGAGAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGG	Qy 111 AAAAGGAGGGGGGGGGGGGGAGAAAAAAAAAAAAA
	QY 218 TSGRGGLGGQGAGAAAAAAAAGGAGGGGGGGGGGGGGG
Oy 358 AAAAAAAGGGGGGGGGGGGGGGGGGGAAAAAAAAAAGGGG	298 GLIGNGGNGGSGGTSVATGGAGNGGAGGAGGAGCAGLIGNGGNGGSGGMGDAP 317 GGAGQGGYGGGQQTSGRGGLGGGG 1

Qy 355 AGA : Db 409 SGA Qy 413 GGL	SRGGLG SDGGAG AAAAAA	oy Oy	254OGTSGRGGLGGOGAGAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGOCAG 306
Db 467 GGA		og .:	
525		<u>3</u> 8	520GUGISGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEG
Qy 524 AAA Db 581 AGD	AAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGACQGGY- 576 	oy G	373 GYGGLGSQGTSGRGGLGGGGAAAAAAAAAAAAGGGGYGGLGSQGTSGRGCLGGGGGG 432 - - - - - - - - - - - - -
Qy 577 -GG 1 Db 640 TGG	-GGLGSOGTSGRGGLGGQCAGAAAAAAAAGA-GQGGYGGL-GSOGTSGRGGLGGQ 631 	Q Q	433 AAAAAAAAAGGGGY-GGLGSQGTSGRGGLGGQGAGAAAAAAAAGGAGGGG 487
Qy 632 GAG 1 Db 698 GGG	GAGAAAAAAAAGAGGGGGGGGGGGGGGGGGGGGGGGGG	Qy Dp	488 YGGLGSQ-GTSGPGGYGPGQQTSGRGGLGGOGAGAAAAAAAAAGGAGOGGYGGLGSQ 544
RESULT 14 H70987 hypothetical N;Alternate n C;Species: My C;Date: 17-Ju	RESULT 14 H70987 hypothetical glycine-rich protein Rv1759c - Mycobacterium tuberculosis (strain H37RV) N.Alternate names: wag22 antigen homolog C.Species: Mycobacterium tuberculosis C.Species: Mycobacterium tuberculosis C.Species: Mycobacterium tuberculosis	Oy Oy Oy	545 GTSGRGGLGGOGAGAAAAAAAGAGAGGCGYGCLGSOGTSGRGCLG 591
C; Accession: R; Cole, S.T.; ; Connor, R.; Rajandream, M	H70987 Brosch R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Davles, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.	Qy Dp	642 AAAGGAGGGYGGLGSQGTSGPGYG 667
A; Authors: Sc A; Title: Deci A; Reference n	Jaf. 244, 1536 gares, K.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. lohering the biology of Mycobacterium tuberculosis from the complete genome number: A70500; MUID:98295987; PMID:9634230		B70812
A; Accession: A; Status: pre A; Molecule ty	H70987 eliminary; nucleic acid sequence not shown; translation not shown fpe: DNA end Ana		thetical glycine-rich protein Rv0834c - Mycobacterium tuberculosis (strain H37RV) ecies: Mycobacterium tuberculosis te: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
A; residues: J A; Cross-refer A; Experimenta C; Genetics: A: Gene: Rv175	1°514 (2002) 1°514 (2002) al source: GB:25890; GB:AL123456; NID:93242245; PIDN:CAB09322.1; PID:92131027 al source: strain H37Rv 59c: wac22		C;Accession: B70Bll2 R;Coll B70Bll2 R;Coll B7.: Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon; Connor, R.; Bovies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rahandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
C; Superfamily: elastin	31.3%; Score 1110.5;	A; Au A; Ti A; Re	thors: Sares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. tle: Deciphering the biology of Mycobacterium tuberculosis from the complete geno ference number: A70500; MUID:98295987; PMID:9634230
Matche	best Local Similarity 38.2%; Fred. No. 1.28-55; Matches 309; Conservative 31; Mismatches 307; Indels 161; Gaps 33;	A; AC A; St A; MO	Accession: B70812 Status: preliminary; nucleic acid sequence not shown; translation not shown Molecule type: DNA
Oy 5 TGG Db 110 TGR	5 TGGQQMGRGSMASGRGGLGGQGAGAAAAAAGAGGQGGYGGL-GSQGTSGRGG 59 	A; Re	
Qy 60 LGG Db 168 AGG	60 LGGGGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGLGGQGAGAAAAA 109 	A; Ge C; Su	834c 1y: elastin
Qy 110 AAA Db 228 GAA	110 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAGG 154 	On Be Wa	Query Match Best Local Similarity 39.9%; Pred. No. 4.1e-54; Matches 303; Conservative 28; Mismatches 306; Indels 122; Gaps 32;
Qy 155 -AG Db 288 MGG		S G	7 GQOMGRGSMASGRGGLGGOGAGAAAAAAAGGAGOGGYGGLGSOGTSGRGGLGGOGAG 66
	205 -AGGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGGGQGGYGGLGS 253 	Qy Dp	67 AAAAAAAAAAGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGAG 118 : :

					_					
175 308	367	275 425	321	376 541	430 594	489 652	528 712	588 768	642 828	
OGGYGGLGSQGTSGRGGLGGQCAGAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGG	-PGQQTSGRGGLGGQGAGAAAAAAAAAGAGGGGYGGLGSQGTSGRG 	GLGGGGAGAAAAAAAAAGAGGGGGGGGGG-SQGTSGRGGLGGGGGAGAAAAAA 	276 AAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQ 	GGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGAGGGGGYGG	LGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGGG : :	AGAAAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGAAAA	TSGATGGAGGDGVFEGIAVLGLGFGGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAA	529 AAGGAGGGGGGGGGGGGGGGGGGGGAGAAAAAAAAAA	GLGGGGAGAAAAAAAAAAGGAGGGGGGGGGGGGGG	AAGGAGQGGYGGLGSQGTS-GPGGYG 667
119	176 309	223 368	276 426	322 485	377 542	431 595	490 653	529 713	589 769	643
çy d	Oy Dp	ç q	Oy Dp	Qy Db	Qy	ογ Op	Ογ Dp	Qy Db	oy Op	Qy

Search completed: December 18, 2002, 16:22:16 Job time : 30.4171 secs

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
```

OM protein - protein search, using sw model

December 18, 2002, 16:06:04; Search time 10.6063 Seconds (without alignments) 2702.183 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-490-291-6 3543 1 MASWIGGOWGRGSWASGRG.....TSGIRIRAPSTSFEHHHHHH 691

Scoring table:

112892 seqs, 41476328 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched: Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_40:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description	P19837 nephila cla	O53553 mycobacteri	9	P46804 nephila cla	006794 mycobacteri	33	_	_	_	Q99372 rattus norv	Q10637 mycobacteri	P10496 phaseolus v		P03211 epstein-bar	-	-		P15502 homo sapien	Q50615 mycobacter1	P33485 pseudorabie	Q50630 mycobacteri	P09789 petunia hyb	S	P30754 riftia pach	m	Q01149 mus musculu	Q10707 mycobacteri	₹	P12105 gallus gall		P02466 rattus norv	8 homo	PO8121 mus musculu
ID	SPD1_NEPCL	YZ08_MYCTU	FBOH_BOMMO	SPD2_NEPCL	WA22_MYCTU	YQ34_MYCTU	Y278_MYCTU	Y747_MYCTU	ELS_MOUSE	ELS_RAT	YD25_MYCTU	GRP2_PHAVU	ELS_BOVIN	EBN1_EBV	CA21_ONCMY	ELS_CHICK	CA21_RANCA	ELS_HUMAN	YI18_MYCTU	VNUA_PRVKA	YP91_MYCTU	GRP1_PETHY	LORI_MOUSE	CAFF_RIFPA	GRP_ARATH	CA21_MOUSE	YK98_MYCTU	YI40_MYCTU	CA13_CHICK	CA13_HUMAN	CA21_RAT	CA12_HUMAN	CA13_MOUSE
DB	-	-	Н	Н	Н	Н	-	Н	Н	-	Н	-		-	-	-	-	-	Н	-	~	Н	П	-	-	-	-	Н	Н	Н	~	~1	
a Query Match Length	747	1901	5263	627	914	778	957	801	860	864	603	465	747	641	1356	750	1355	730	498	1733	543	384	481	1027	338	1372	491	515	1262	1466	1372	ч.	1464
8 Query Match	60.3	36.2		ď.	31.3	•		27.9			24.5		21.9		21.1		•	19.5	19.4		19.2				18.2	•	17.7	7.	7.	17.6	7.	17.2	17.1
Score	2138	1281.5	1174.5	1138.5	1110.5	1049.5	1032	988.5	916	939	869	853.5	777.5	763	747.5	709	693.5	691	989	681	678.5	673	699	656.5	644	630	929	626	624.5	622	608	608	607
Result No.	1	2	Э	4	Ŋ	9	7	80	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	36	27	28	29	30	31	32	33

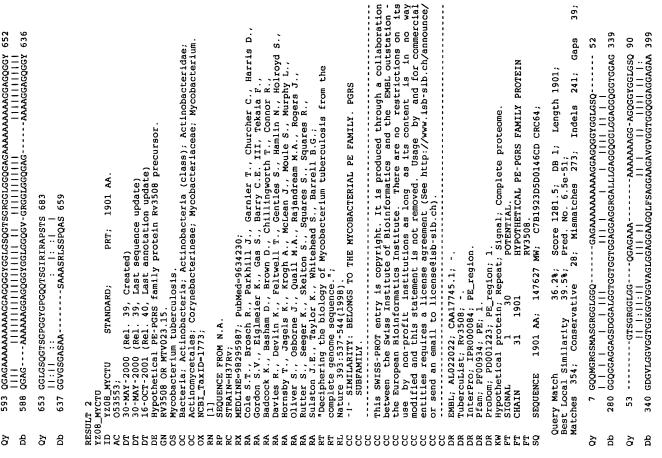
		P28481 mus musculu P02452 homo saplen O46392 canis famil		
CA11_CHICK CA13_BOVIN	YA68_MYCTU CA11_CANFA CA21_HUMAN	CA12_MOUSE CA11_HUMAN CA21_CANFA	CA21_BOVIN CA11_MOUSE CA21_CHICK	CA24_CAEEL
				н
1453 1049	463 1460 1366	1459 1464 1366	1364 1453 1362	1758
17.1 17.0	17.0 17.0 17.0	16.9 16.9 16.9	16.8 16.7 16.2	15.9
606.5	602.5 602.5 602	599.5 599.5 599	595 592.5 573.5	563
34 35	36 37 38	39 41 41	4 4 4 4 3 6	45

ALIGNMENTS

```
QGAGAAAAAAAAAAGGAGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                      30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv3508 precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         36.2%; Score 1281.5; DB 1; 39.5%; Pred. No. 6.5e-51; Live 28; Mismatches 273;
                                                                                                                                                                              1901 AA
                                                                           683
                                                                                                   637 GGVGSGASAA-----SAAASRLSSPQAS
                                                                         GGLGSQGTSGPGGYGPGQQTSGIRIRAPSTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RV3508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1901 AA; 147627 MW;
                                                                                                                                                                                                          30-MAY-2000 (Rel. 39, Created)
                                                                                                                                                                                                                                                                  RV3508 OR MTV023.15.
Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 39.5
Matches 354; Conservative
                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1901
                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                         STRAIN-H37RV
                                                                                                                                                                           YZ08_MYCTU
O53553;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GNAL
                593
                                                                         653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      91
                                            셤
                                                                                                   g
                                                                         ð
                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                  35;
                                                                                                                                                                                                                                                                                V -> L (IN REF. 1).

S -> T (IN REF. 1).

NPGLSGCDVLIQALLEVVSALQLIGSSSTGQVNYGSAGQA
TQIVGQSVYQAL -> ILVFLDVMSSFKLFSRLFLLLSRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GCYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGA 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA-------AGGAGQGGYGGLG 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----SGP 499
                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                                                                                                                                                           25 QGAGAAAAAAAAAGGAGGGYGGLGSQGT--SGRGGLGGQGAG-AAAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGGAGQGGYGGLGSQG-AGRGGLGGQGAGAAAAAGGAGQGGYGGLGGQGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SQCTSGRGGLGGQGAGAAAAAA--------AAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGQGAGAAAA-------AAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGYGP-GQQTSGRGGLGGQGAGAAAAAAA-----AGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------AAAAAGGAGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                 49; Indels 184;
                                                                                                                                                                                                                                                                                                                                                                                   Length 747;
                                                                                                                                                                                                                                                                                                                                                      850E44B0D649E012 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                     8; DB 1;
4.5e-89;
                                                                                                                                                                                                                                                                                                                                                                                   ; Score 2138; Dl
; Pred. No. 4.5e.
20; Mismatches
                                                                                                                                                                                                                                                                                                                                          (IN REF.
                                                                                                                                                                                                                                                                                                                                                        60528 MW;
                                                                                                                                                                                                                                                                                                                                                                                     60.38;
66.38;
                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 66.3
Matches 498; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGQGAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                      747 AA;
1159
22022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
2302
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
2302
23022
23022
23022
23022
23022
23022
23022
23022
23022
23022
230
                                                                                                                                                                                         REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                          REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          476
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
```



m

```
.089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        168
                                                                                                                           578
                                                                                                                                                      259
                                                                                                                                                                                 636
                                                                                                                                                                                                                                                                                            756
                                                                                                                                                                                                                                                                                                                       393
                                                                                                                                                                                                                                                                                                                                                   816
                                                                                                                                                                                                                                                                                                                                                                               435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SINANGGAGGUGGTGGKGGAGGAGTLGVGGSG--GTGGDGGDAGSGGGGGFGGAAGKAG 993
             GADAPASTGLTGGTGFAGGAGGVGGQGGNAIAGGINGSGGAGGTGGQGGAGGMGGSGADN 459
                                                                                               207
                                                                                                                                                                                                            311
                                                                                                                                                                                                                                      969
                                                                                                                                                                                                                                                                                                                                                                                                                                   -----AAAAAAAGGAGGGYGGLGSQG----TSGRGGLGGQGA----GAAAAA 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                             877 NGGNAGVGLTAKAGDGGAAGNGGNGGAGGAGGAGDNNFNGGQGGAGGQGGQGGLGGASTT 936
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hexapoda;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda.
Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Bombycoidea; Bombycidae; Bombyx.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1114 NFNGGOGGAGGOGGOGGLGGASTISINANGGAGGNGGTGGKGGAGGAGTLGVGGSG 1169
                                                                                                               DATATGATGGTGFSGGAGGAGGAGGNTGVGGTNGSGGQG-GTGGAGGAGGAGGAGGADNPT
                                                                                                                                                                      ------GGLGGQG--AGAAAAAAAAAGG-------AGQGGYGGLGSQGT
                                                           ----SGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQ----
                                                                                                                                                     --GGYGGLGSQGTSGRGGLGGQG--AGAAAAAAAAGG-AGQGGYGGLGSQGT---SGR
                                                                                                                                                                                                        GGLGGQCACAAAAAAAAAAGGAGQGGYGG------LGSQGTSGRGGLGGQ-GAGAAAAA
                                                                                                                                                                                                                                                                312 AAAAAGGA----GQGGYGGLGSQ----GTSGPGGYGPGQQTSGRGGLGGQGA-----
                                                                                                                                                                                                                                                                                          697 GSSATGGAGFAGGAGGEGGAGGNSGVGGTNGSGGAGGAGGKGGTGGAGGSGADNPTGAGF
                                                                                                                                                                                                                                                                                                                       ----GAAAAAAAAAAGGA-GQGGYGGL-----GSQGTSGRGCLGGQGA-----
                                                                                                                                                                                                                                                                                                                                                757 AGGAGGTGGAAGAGGAGGATGTGGTGGVVGATGSAGIGGAGGRGGDGGDGASGLGLGLSG
                                                                                                                                                                                                                                                                                                                                                                             476 AAAAAGGA----GQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GAGQGGYGGLGSQCTS-----GRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --TSGRGGLGGQGA----GAAAAAAAAAGGA----GQGGYGGLGSQGTSGPGGYG 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Esnault C.,
Li Z.-G.;
gene.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-20330362; PubMed-10871375;
Zhou C.-Z., Confalonieri F., Medina N., Zivanovic Y.,
Yang T., Jacquet M., Janin J., Duquet M., Perasso R.,
"Fine organization of Bombyx mori fibroin heavy chain
Nucleic Adids Res. 28:2413-2419(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FBOH_BOMMO STANDARD; PRT; 5263 AA. P05790; Q26379; Q17220; Q100v-1988 (Rel. 09, Created) 16-00T-2001 (Rel. 40, Last sequence update) 16-00T-2001 (Rel. 40, Last annotation update) Fibroin heavy chain precursor (Fib-H) (H-fibroin).
Bombyx mori (Silk moth)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=7091;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FBOH_BOMMO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FIBH.
              400
                                                                                                                         520
                                                                                                                                                    208
                                                                                                                                                                                                         260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    581
                                                                                              169
                                                                                                                                                                                                                                                                                                                       356
                                                                                                                                                                                                                                                                                                                                                                            394
                                                                                                                                                                                                                                                                                                                                                                                                                                   436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                               g
                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                              à
                                                                                                                     8
                                                                                                                                           à
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                         ç
                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to licenseelib-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RA Tanaka K., Kajiyama N., Ishikura K., Waga S., Kikuchi A., Ohtomo K., Takadi T., Mizuno S.;
Inght chains of the site of disulfide linkage between heavy and Inght chains of silk fibroin produced by Bombyx mori.";
Elochim. Blophys. acta 1421:92-103(1999).

INSOLUBLE AND CHEMICALLY INERT FIBRE.
INSOLUBLE AND CHEMICALLY INERT FIBRE.

INSOLUBLE AND CHEMICALLY INERT FIBRE.

INSOLUBLE AND CHEMICALLY INERT FIBRE.

INSOLUBLE BODD. HEAVY-LIGHT CHAIN ASSEMBLY IS ESSENTAL FOR THE EFFICIENT INTRACELLULAR TRANSPORT AND SECRETION OF FIBROIN.

I SUBJUCTEDE BOND. HEAVY-LIGHT CHAIN ASSEMBLY IS ESSENTAL FOR THE EFFICIENT INTRACELLULAR TRANSPORT AND SECRETION OF FIBROIN.

I SECTION OF SILK GLANDS.

I SECTION OF SILK GLANDS.

I DOMPOSED OF ANTIPARALLEL BETA SHEETS. THE STRANDS OF THE BERA SHEETS RUN PARALLEL TO THE FIBER AXIS. LONG STRETCHES OF SILK CHAN-SHEADS OF MICROCRYSTALLINE ARRAYS OF (-GLY-SER-GLY-ALA-GLY-ALA-)N INTERRUPTED BY REGIONS CONTAINING BULKERS TERMING WITH THE FIBER IS COMPOSED OF MICROCRYSTALLINE ARRAYS ALTERNATING WITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mita K., Ichimura S., James T.C.; "Highly repetitive structure and its organization of the silk fibroin
SEQUENCE OF 1-168 FROM N.A.
SEQUENCE OF 1-168 FROM N.A.
MEDLINE-80045039; PubMed-498286;
Trujimoto Y., Suzuki Y.;
"The DNA sequence of Bombyx mori fibroin gene including the 5'
"The DNA sequence of Bombyx mori fibroin gene including the 5'
flanking, mRNA coding, entire intervening and fibroin protein coding
                                                                                                                                                                                                                                                                                                                                                                                                   on the secondary
                                                                                                                                                                             PARTIAL SEQUENCE FROM N.A. MEDIJABE-7921121; PubMed-455439; Tsujimoto Y., Suzuki Y.; Structural analysis of the fibroin gene at the 5' end and its surrounding regions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 5179-5263 FROM N.A., AND DISULFIDE BONDS
                                                                                                                                                                                                                                                                                                                 PARTIAL SEQUENCE FROM N.A.
STRAIN-Kinshu X Showa;
MEDLINE-89094868; PubMed-3210244;
MILA K., Ichimura S., Zama M., James T.C.;
"Specific codon usage pattern and its implications structure of silk fibroin mRNA.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
FIBROIN HEAVY CHAIN.
HIGHLY REPETITIVE.
                                                                                                                                                                                                                                                                                                                                                                                                     structure of silk fibroin mRNA.";
J. Mol. Biol. 203:917-925(1988).
[5]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PARTIAL SEQUENCE FROM N.A. MEDLINE=94365842; PubMed=7916056;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          J. Mol. Evol. 38:583-592(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; V00094; CAA23432.1; EMBL; V00094; CAA23433.1; EMBL; X1869; CAA32076.1; EMBL; X1869; CAA32076.1; EMBL; M3578; AAA27839.1; EMBL; M5578; AAA27839.1; PTR; S01844; S01844. S01844. S11k; S19nal; Repeat. 2 S167AL 149 5206 HI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AF226688; AAF76983.1; -.
                                                                                                                                                                                                                                                                               Cell 16:425-436(1979).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-J-139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
```

```
134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 188
  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                        27;
                                                                                                                        4456
                                                                                                                                                       4516
                                                                                                                                                                                                                      4636
                                                                                                                                                                                                                                                     4696
                                                                                                                                                                                                                                                                                                                                                                                                                  4995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5115
                                                                                4697 GSGSGAGAGSGSGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGAGYGA 4756
                                                                                                                                                                                                                                                                                                            5055
                                                                                                                                        170
                                                                                                                                                                                                                                                                                                                                    443
                                                                                                                                                                                                                                                                                                                                                                   491
                                                                                                                                                                                                                                                                                                                                                                                 340
                                                                                                                                                                       207
                                                        Gaps
                                                                                                       GGQGAGAAAAAAAAAGGAGQG-----GYG---GLGSQGTSGRGGLGGQGAGAAAAAA
                                                                                                                                                   GAGAAAAAAAAAAGGGGGGGGGGGGGGGGGGGG-----GLG-GQGAGAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                              GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQG-------GYG-GL
                                                                       2 ASMTGGQQMGRGS-MASGRGGLGGQGAGAAAAAAAAAAAAGGAGGGGYGGLGSQGTSGRGGL
                                                                                                                      AAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAGGAGQG-GYG-GLGSQGTSG
                                                                                                                                                                     PG---GYGPGQQTSG-----RGGLG------GQGAGAAAAAAAAGGAG
                                                                                                                                                                                   4517 ACACACACACACACGSGFGPYVANGCYSCYEYAWSSESDFGTGSGAGAGSGAGAGSGAGAGY
                                                                                                                                                                                                                                                                                                                                                                                                                                  RGGLGGQGAGAAAAAAAAAAGGAGQG-GYG-GLGSQGTSGRGGLGGQGAGAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                GGAGQG------GYG-GLGSQGTSGRGGLGGQCAGAAAAAAAAAAAAGGAGQG---GYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 G-----GYG-----GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                   -----GLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPG-GYGP
                                                                                                                                                                                                                                                                                                   GQQTSGRGGLG------GQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                  GSQCTSGPG---GYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGGGGYGGLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                 GAGYGAGAGYGAGYGAGAGSGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAG
                                                3e-46;
ss 305; Indels 131;
                                        DB 1; Length 5263;
INTERCHAIN (WITH LIGHT CHAIN)
                -> V (IN REF. 2).
; 8EE11D3A0A47440E CRC64;
                                                                                                                                                                                                                                    254 QGTSGRGGLGGQGAGAAAAAAAAAGGAGQG-------GYG-
                                      Score 1174.5; D
Pred. No. 6.3e-4
19; Mismatches 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        627 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5116 AGAGSGAASGAGAGAGAGTGSSG 5138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -GLGSQGTSGPG-GYGPGQQTSG 674
                10 10 C -
5263 AA; 391586 MW;
                                                       49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                       33.18;
                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Rel. 32, (Rel. 32, (Rel. 41,
                                                Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1995
01-NOV-1995
15-JUN-2002
                                                 Best Local Sim
Matches 318;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SPD2_NEPCL
P46804;
                CONFLICT
DISULFID
                                       Query Match
                                                                                                                      4397
                                                                                                                                                                                                                                                                                                   341
                                                                                                                                                                                                                                                                                                                                   392
                                                                                                                                                                                                                                                                                                                                                                                 4877
                                                                                                                                                                                                                                                                                                                                                                                                                 4937 (
                                                                                                                                                     4457
                                                                                                                                                                                                                                                                                                                   1757
                                                                                                       61
                                                                                                                                      113
                                                                                                                                                                     171
                                                                                                                                                                                                     208
                                                                                                                                                                                                                                                                    288
                                                                                                                                                                                                                                                                                                                                                                                                                                  549
                                                                                                                                                                                                                                                                                                                                                                   444
                                                                                                                                                                                                                                                                                                                                                                                                 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 607
FFFS
                                                                                                                     셤
                                                                                                                                      à
                                                                                                                                                    g
                                                                                                                                                                    ð
                                                                                                                                                                                     g
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                               음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25555
                                                                       ð
                                                                                      g
                                                                                                       õ
                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ä
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lab-sib.ch).
                                                                                                             187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         79 AGQGGYGGLG-SQGTSCRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQ----GTSGR 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 GGYGPGQ------QGPGGYGP-GQQGPSGPG-----SAAAAAAAAAAGPGGYG 42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -GQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG----GLGGQGAGAAAAAAAAGG
                                            Nephila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15 APPROXIMATE TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         190; Indels
             Nephila clavipes (Orb spider).
Eukaryota: Metazoa; Arthropoda; Chelicerata; Arachnida;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CB9B63779B2C594B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 1138.5; DB 1;
Pred. No. 7.6e-45;
Spidroin 2 (Dragline silk fibroin 2) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mismatches
                                                                                                  MEDLINE=92406876; PubMed=1527052;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54184 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M92913; AAA29381.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        47.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   464
488
515
530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 516
627 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                    FROM N.A.
                                                       NCBI_TaxID=6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         325;
                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         243
```

36;

S

us-09-490-291-6.rsp

```
STRAIN=CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haff D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                          470
                                                                                                                                                                                                                                                                                                                 590
                                                                                                       412
                                  357
                                                                                                                                                                                                                                                                                                                                                                                                      s:
                                                                                                                                                                                                        GP-GQQGPSGPG-----SASAAAAAAAAAAGPGGYGP-GQQGPGGYAPGQQGPSGPG 417
                                                                                                                                                                                                                                                                         SASAAAAAAA -- AGPGGYG ---- PGQQGPGGYAPGQQ ----- GPSGPGSAAAAAA - 463
                                                                                                                                                                                                                                                                                                                                                                                  ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDINE-9825987; PubMed=9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris I Gordon S.V., Eiglameier K., Gas S., Barry C.E. III. Tekaia F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamin N., Holroyd S. Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the Complete genome sequence.";
Nature 393:537-544(1998).
                                  GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGA
                                                  QGPGGYG----PGQQGPGGYGPGQQGLSGPGSAAAAAAAAAGPGQQGPGGYGP-GQQGPSGP
                                                                                                 311 AAAAAAAGPGQQGLGGYGP-GQQGPGGYGPGGYGPGGYGPGSASAAAAAGPGQQGPGGY
                                                                                                                                                                        --AGPGGYGP-AQQGPSGPG------IAASAASAGPGGYG-----PAQQGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Whole genome comparison of Mycobacterium tuberculosis clinical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-MAY-2000 (Rel. 39, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       GQGGYGGLGSQGTSGPGGYGPGQQTSGIRIRA 679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SSAALSSVISNAVSQIGASNPGLSGCDVLIQA 587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WAG22 antigen precursor.
WAG22 OR RV1759C OR WT1807 OR MTCY28.25C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-MAY-2000 (Rel. 39, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WA22_MYCTU
006794;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WA22_MYCTU
                                                                                                                                                                        413
                                                                                                                                                                                                                                                                                                              531
                                                                                                                                                                                                                                                                                                                                                464
                                                                                                                                                                                                                                                                                                                                                                                  591
                                                             g
                                                                                                                                    a
                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                      å
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                 ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                      33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GAAGAGGAGGFSFSTAGGAGGAGGAFTTGGVGGAGGQGHTGGAGGAGGAFGAGG 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -AGQGGYGGLGSQGT---SGRGGLGGQGAGAAAAAAAAAAGGAG-----QGGYGGLGS-- 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  348 AGGAGCTGGAG--GTVFGGGKGGAGGAGGNAGMLFGSGGGGGTGGFGFAAGGQGGVGGSA 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---QGTSGRGGLGGQG--AGAAAAAAAAAGGAGGGYGGLGSQG--TSGRGGLGGQGAG 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GYĠGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAG 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TGRPLIGNG--ANGAPGTGANGGDAGWLIGNGGAGGSGAKGANGGAGGPGGAAGLFGNGG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LGGQGAGAAAAAAAAAGGA-----GQGGYGGLGSQGTS---GRGGLGGQGAGAAAAA 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AGGAGGTATANNGIGGAGGAGGSAMLFGAGGAGGAGGAATSLVGGIGGTGGTGGNAGMLA 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAAAAGGAG-----QGGYGGLGSQG----TSGRGGLGGQG-----AGAAAAAAAAAGG 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LIGRPLIGNGDSGTPGTGDDGGAGGWLFGNGGNGGAGA-AGINGSAGGAGGAGGILFGTG 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   585 GAGGAGGVGTAGAGGAGGAGGSAFL-----IGSGGTGGVGGAATT-TGGVGGAGGN 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                    TGGQQMGRGSMASGRGGLGGQGAGAAAAAAAAAGGAG----QGGYGGL-GSQGTSGRGG 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -AGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAAGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AGLLIGAAGLGGCGGGAFTAGVTTGGAGGTGGAAGLFANGGAGGAGGTGSTAGGAGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        488 YGGLGSQ-GTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAG--QGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A-----AAAAAAAAAAGGGGYGGL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVLIGNGGEGGIGALAGKSGFGGFGGLLLGADGYNAPESTSPWHNLQQDILSFINEPTEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31.3%; Score 1110.5; DB 1; Length 914; 38.2%; Pred. No. 1.6e-43; Live 31; Mismatches 307; Indels 161;
                                                                                                                                                                                                                                                                         Antigen; Repeat; Signal; Complete proteome.
SIGNAL 1 30 POTENTAL.
SEQUENCE 914 AA; 74354 MW; F6953C3DBBBE6ACB CRC64;
                                                                                                                                                                      EMBL; 295890; CAB09322.1; -- EMBL; AE007040; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                      TIGR; MT1807;
Tuberculist; Rv1759c;
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE; 1.
FRAMESHIFT IN POSITION 85.
                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                  Matches 309;
                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      635
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
```

```
Best Local Similarity
                    Matches
                                                                             129
                                                                                                           28
                                                                                                                                         189
                                                                                                                                                                                                                                  176
                                                                                                                                                                                                                                                              300
                                                                                                                                                                                                                                                                                            236
                                                                                                                                                                                                                                                                                                                                                                                                                 343
                                                                                                                                                                                                                                                                                                                                                                                                                                               456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        632
                                                                                                                                                                                                                                                                                                                                                      284
                                                                             g
                                                                                                           ò
                                                                                                                                      a
                                                                                                                                                                       ö
                                                                                                                                                                                                셤
                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [2]
SEQUENCE FROM N.A.
SEQUENCE 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Fleischmann R.D., Hockson R., Gwinn M.L., Haft D., Hickey E.,
Foeterson J., DoBoy R. Unayam L.A., Ermolaeva M.D., Salzberg S.L.,
Kolonay J.F., Nelson W.C., Unayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
    641
                                                                                                                                                                                                                                                                                                                                                                                                                                               .
S
                                                                                                                                                                                                                                                                                       Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                           Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamin N., Holroyd Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Seeger K., Skelton S., Squares R., Sulston J.E., Taylor K., Whitehead S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

"Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
GQG-----AGAAAAAAAAAAGAGQGGYGG-LGSQGTSGRGGLGGQGAGAAAAAA
                      laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           V -> L (IN REF. 2).
Q -> H (IN REF. 2).
A -> T (IN REF. 2).
DAB2OFE58E4999E7 CRC64;
                                                                                                                                                                                            01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
19potchetical PE-PGRS family protein Rv2634c.
RV2634C OR WT2712 OR MTCY41.04C.
                                                                                                                                                                     778 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Complete proteome
                                                             642 AAAGGAGQGGYGGLGSQ--GTSGPGGYG 667
                                                                               EMBL; 280225; CAB02341.1; -. EMBL; AE007103; AAK47026.1; ALT_INIT
                                                                                                                                                                                                                                                                                                                                                                  STRAIN=H37Rv;
MEDLINE=98295987; PubMed=9634230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tuberculist; Rv2634c; -.
Tuberculist; Rv2634c; -.
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE; 1.
ProDom; PD00123; PE_region; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         63131 MW;
                                                                                                                                                                                                                                                                           Mycobacterium tuberculosis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nature 393:537-544(1998).
                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hypothetical protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         778 AA;
                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBFAMILY.
                                                                                                                                                                 YQ34_MYCTU
P71933;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                      RESULT 6
YQ34_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TIGE;
à
                              셤
                                                             ò
                                                                                      g
```

Length 778;

DB 1;

Score 1049.5;

29.68;

Query Match

```
34;
                                                                                                         GAPGGPGGLLLGNGGNGGSGAPGQPGGAGGDAGLIGNGGTGGKGGDGLVGSGAAGGVGGR 188
                                                                                                                                                     115
                                                                                                                                                                                               243
                                                                                                                                                                                                                                                                                                                                                                  344
                                                                                                                                                                                                                                                                                                                                                                                                            283
                                                                                                                                                                                                                                                                                                                                                                                                                                                    404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    678 AG--GIGGIGGAGGIIGGSGGAGGIIGWAGAAGGIGAGGIGGGGGIGGGGGG-GNGGNGGI 733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Harris D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .
S
                            Gaps
                                                                    57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-98295987; PubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris
Gordon S.V., Eiglameler K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davles R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd
                                                                  GQQMGRGSMASGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQ------GTSGR
                                                                                                                                                   GG--LGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAG
                                                                                                                                                                                                                                       116 GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYG
                                                                                                                                                                                                                                                               244 VGGSGGVGGV--FGNGGFGGAGGLGAAGGVGGAASYFCTGGGGGGVGGDGAPG--GDGGAG
                                                                                                                                                                                                                                                                                                                        PGQQTSGRGGLGGQGAGAAAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                     P--LLIGNGGVGGLG-GAGAA------GNGGAGGM----LLGDGGAGGQGGPAVAGV
                                                                                                                                                                                                                                                                                                                                                                                                          AAAAAGGAGQG-----GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAG-----GAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                  345 LGGMPGAGGNGGNANWFGSGGAGGQGGTGLAGTNGVNPGSIANPNTGANGTDNSGNGNQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GGYGGLGSQGTSGR-GGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---GAGGAGGIGETDGSAGGVATGGEGGDGATGGVDG-GVGGAGGKGGQGHNTCVGDAFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             579 LGSQGTSGRGGLGGQ--GAGAAA----AAAAAAAGGGGGYGGLGSQGTSGRGGLGGQ
                                                                                                                                                                                               GGWLLGNGGTGGAGGAAGATLVG-GTGGVG--GATGLIGSGGFG--GAGGAAAGVGTTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             512 GDGGIGGDGNGALGAAG--GNGGTGGAGGNGGRGGMLIGNGGAGGAGGAGGTGGTGGGGAAGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      521 AAAAAAAAAAAGGA-GQGGYGGLGSQGTSGRGGLGGQ-GAGAAAAAAAAAAAAAGGGGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              403 AAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGA-GQGGYGGLGSQGTSG-
                            66;
                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 734 GATGGGGGDFALGGNGGAGGAGGSPGGSSGIQGNMGP----PGTQ 774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last amontation update)
Hypothetical PE-PGRS family protein Rv0278c precursor.
RV0278C OR MT0291 OR MTV035.06C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GAGAAAAAAAAAGGAGGGYGG-----LGSQGTSGPGGYGPGQQ
¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          957
                        291; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-H37RV;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X278_MYCTU
P56877;
```

~

```
05381
                                                                                                                                                                      408
                                                                                                                                                                                                                  296
                                                                                                                                                                                                                                                             162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                             셤
                                                                          å
                                                                                                               셤
                                                                                                                                                                 ò
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isp-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30;
                                                                                                                                                                                                                                     Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Bobby R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGIG--GAGGAGGNAGLLFGSGGSGGAGGFGF-ADGGQGGPGGNAGTVFGSGGAGGNGGV 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----GAGA .357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                       "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GOOMGRGSMASGRGGLGGOGAGAAAAAAAAGGAGOGGYGGL-GSOGTSGRGGLGGOGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           126 GANGGDGGWLIGNGGAGGSGAAGVNGGAGGNGGAGGNGGAGGLIGNGGAGGAGGVASSGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GAAAAAAAAA--GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    186 GGSGGAGGNAMLFGAGGAGGAGGGVVALTGGAGGAGGAGGNAGLLFGAAGVGGAGGFTNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --QGGYGGLGSQG----TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                246 SALGGAGGAGGLFATGGVGGSGGAGSSGGAGGAGGAGGLFGAGGTGGHGGFADSSFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              173 GYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGG------QGGYGGLGSQ-GTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       222 GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   282 GQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG----LGSQGTSGPGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GOGFAGGIG -- GAGGTPGLIGNGGNGGASAVTG -- GNGGIGGTGVLIGNGGNGGSGG
       ŗ.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HYPOTHETICAL PE-PGRS FAMILY PROTEIN
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 tch 29.1%; Score 1032; DB 1; Length 957; al Similarity 37.4%; Pred. No. 5e-40; 292; Conservative 29; Mismatches 313; Indels 146;
                                                                                                                                                                                                                                                                                                                                                                        laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
--- SIMILARITY: BELONGS TO THE MICOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RV0278C.

M -> I (IN REF. 2).

MISSING (IN REF. 2).

R -> G (IN REF. 2).

M; 71EBABD417FBA47C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; Arrows 1.

Pfam; PF00934; PE; 1.

Hypothetical protein; Repeat; Signal; Complete proteome.

### Prothetical Protein; Pepeat; Signal; Complete Proteome.

### Prothetical Protein; Pepeat Proteome.

### Prothetical Protein 
                                                                                                                                                                                                           STRAIN-CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; AL021930; CAA17353.1; -.
EMBL; AE006936; AAK44511.1; ALT_INIT.
TIGR; MT0291; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000084; PE_region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81905 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               338 YGPGQQTSGRGGLGGQ-----
                                                                                                                                         Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tuberculist; Rv0278c;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  40
158
807
957 AA;
                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                               Bishai W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99
  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cole S.T., Brosch R.J., Gars S., Barry C.E. III. Tekala F.,
Gordon S.V., Eiglmeler K., Gas S., Barry C.E. III. Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quall M., Rajandream M.A., Ragers J.,
Rutter S., Seeger K., Skelton S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
Deciphering the biology of Mycobacterium tuberculosis from the
Complete genome sequence.";
Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S.L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             830
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   713 GNGGLFGTGGTGGAGTNFGAGGNGGNGGLFGAG--GTGGAAGSGGSGITTGGGGHGGNAG 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 653 GGAGGRGGDAGLLFGAAGTGGHGAAGGVPAGVGGAGGNGGLFANGGAGGAGGFNAAGGNG 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Dodson R., Gwinn M.L., Haff D., Hickey E., Kolonay J.F., Nelson W.C., Umayam A.A., Ermolaeva M.D., Salzberg S.L. Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  479 IGAGK--AGVGGVSGLLLGLDGFNAPASTSPLHTLQQNVLNVVNEPFQTLTGRPLIGNGA
                                                                                                                                                         NGTPGTGADGGAG-GWLFGNGANGTPGTGAAGGAGGWLFGNGGNGGHGATNTAATATGGA
                                                                                                                                                                                                                                          GQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAAGGA----GQGGYGGLG--SQGTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         771 LLSLGASGGAGGSGCASSLAGGAGGTGCNGALLFGFRGAGGAGGHGGAALTSIQQGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      609 AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQ-GTSGPGGYG
                                                                                --GGAGGI-LFGTGGNGGTGGIATGAGGIGGAGGAGGVSLLIGSGGTGGNGGNSIGVAGI
                                                                                                                                                                                                                                                                                                                                                                                            GGLGGQGA-----GAAAAAAAAAAAAGG--AGQGGYGGLG----SQGTSGPGGYGPGQQTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      559 AAAAAAAAAGGAG-----QGGYGGLGSQGT----SGRGGLGGQGAGAAAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  30-NAY-2000 (Rel. 39, Created)
30-NAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv0747 precursor.
RV0747 OR MT0772.5 OR MTV041.21.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 801 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-98295987; PubMed-9634230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
STRAIN-CDC 1551 / Oshkosh;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bishai W.
```

602 AAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA-AAAAAAAGGAGQGGYGG--LGSQ

```
28;
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                       287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      625
                                                                                                                                                                                                                                                                                                                                                                 GRPLIGNGANGAPGTGANGGPGGWLIGNGGAGGSGAPGAGAGGNGGAGGLFGSG-GAGGA 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        601
                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                            11 GRGSMASGRGGLGGQGAGAAAAAAAAAAAAAAAAAAAAAAAA 70
                                                                                                                                                                                                                                                                                                                                                                                                       STDVAGGAGGAGGAGGNAGMLFGAAGVGGVGGFSNG-GATGGAGGAGGAG-GLFGAGRER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  288 GGAG-GISKSVGDSAAGGAG-GAPGLIGNGGNGGNGGASTGGGDGGPGGAGGTGVLIGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GTSGRGGLGGQGAGAAAAAAAAAAGAGGGGYGGLG-SQGTSGPGGY-GPGQQTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                         228 GSGGSGNLTGGAGGAGGAGGTLATGDGGAGGTGGASRSGGFGGAGGAGGDAGMFFGSGGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL---GGQGAGAAAAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    319 AGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGA-----GAAAAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGTGGSGGPGATGLGGIGGAGGAALLFGSGGAGGSGGAGAVGGNGGAGGNAGALLGAAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGG----AGQGGYGGLGSQ-GTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----GAAAAAAAAAAAGGAGG--GYGGLGSQGTSGRGGLGGQGAGAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGAGGAGAVG - GNGGAGGNG - GLFANGGAGGPGGFGSPAGAGGIGGAGGNGGLFGAGGT
                                                                                                                                              HYPOTHETICAL PE-PGRS FAMILY PROTEIN RV0747.
                                                                                                                                                                                                                                                                                                                      Indels 107;
                                                                                                                                                                                                                                                                                                27.9%; Score 988.5; DB 1; Length 801;
                                                                                                                           Complete proteome
                                                                                                                                                                                                                                                                         EA54C9BF45A00F41 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GQGGYGGLGSQGTSGRGGLGGQGA-------
                                                                                                                                                                                                                                                                                                         38.3%; Pred. No. 3.05;
tive 35; Mismatches 305;
                                                                                                                                                                                                                                         66
                                                                                                                                                                                                       REF.
REF.
REF.
                                                                                                                                                                                                                                                              (IN REF.
                                                                                                                                                                    G -> S (IN REF.
R -> G (IN REF.
K -> R (IN REF.
S -> G (IN REF.
T -> I (IN REF.
T -> P (IN REF.
T -> A (IN REF.
T -> A (IN REF.
MISSING (IN REF.
                                                                                                                           Signal;
                                                     EMBL; AL021958; CAA17514.1; -.
EMBL; AE006968; AAK45011.1; -.
TIGR; MT0772.5; -.
Tuberculist: Rv0747; -.
InterPro; IPR000084; PE_region.
Pfam; Pf00934; PE; 1.
Hypothetical protein; Repeat; Sig
                                                                                                                                                                                                                                                                         65407 MW;
                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                       295
300
338
                                                                                                                                                                      188
225
227
227
300
338
377
577
580
                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                    CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                   CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           244
                                                                                                                                                                                                                                                                                                                                                                                       71
                                                                                                                                                                                                                                                                                                                                                                                                              170
                                                                                                                                                                                                                                                                                                                                                                                                                                   129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            681
õ
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40;
                                                                                                                                                                                                                                                                                                                                                                                                     184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LGSQGTSGRGGL-GGQGAGAAAAAAAAAAAAAAAAAAAAAAAAA 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene to
SHIGGEGEGE SAVLIGNGGNGGNGGNSGNAGKSGGAPGPSGAGGLLLGEN
                                                                                                                                                                                                                                                                                                                      STRAIN-BALB/C; TISSUE-Lung;
MEDLINE-95130065; PubMed-7829060;
Wydner K.S., Sechler J.L., Boyd C.D., Passmore H.C.;
Wydner K.S., sechler J.L., Boyd C.D., Passmore H.C.;
"Use of an intron polymorphism to localize the tropoelastin gene t mouse chromosome 5 in a region of linkage conservation with human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    108 AAAAAAAG-GAG-QGGYGGLGSQGTSGRGGLGGQGA------GAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               185 VVPQVGAGIGAGGKPGKVPGVGLPGVYPGGVLPGTGARFPGVGVLPGVPTGTGVKAKAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152 AGGAGOG--GYGGLGSQ-----GTSGPGGYG----PGQQTSGRGGLGGQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           39.3%; Pred. No. 1.4e-37;
.ive 36; Mismatches 249; Indels 214;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27.5%; Score 976; DB 1; Length 860;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRINTS; PR01500; TROPOELÁSTIN.
Structural protein; Repeat; Signal; Connective tissue.
SIGNAL 1 27 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OCOBESAAE1EDD7F1 CRC64;
                                                                                                                                                                                (Rel. 34, Last sequence update)
(Rel. 34, Last annotation update)
                                                                                                                                       Ą.
                                                                                                                                       860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR003979; tropoelastin.
                                                                                                                                                                                                       Elastin precursor (Tropoelastin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    860 EL. 71955 MW;
                                                                                                                                                                01-OCT-1996 (Rel. 34, Created)
                                                                                                                                                                                                                                                  Chordata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; U08210; AAA80155.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 39.3
Matches 323; Conservative
                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                     musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    860 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                              NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MGD; MGI:95317;
                                                    | :|
GLNG 799
                                      662
                                                                                                                                                                          01-OCT-1996
01-OCT-1996
                                    GTSG
                                                                                                                                     ELS_MOUSE
P54320;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
        738
                                                               964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        138
                                    629
        셤
                                    à
                                                             a
                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
```

us-09-490-291-6.rsp

```
478
                                                                                                                                                                                                                                                                                                                                     648
                                                                                                                                                                                                                                                                                                                                                                                                                 574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               810
                                                                 GTGVGSQAAAAAAKAAKYGAGGAGVLPGVGGGGIPGGAGAIPGIGGIAGAGTPAAAAAK 364
                                                                                                                       424
                                                                                                                                                321
                                                                                                                                                             371 QGGY-GGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQ 429
                                                                                                                                                                                                                                                                                                                                                                                     649 GAGVPGFGAGAVPGSLAASKAAKYGAAGGLGGPGGLGGPGGLG-----GPGGLGGAGVP 702
                                                                                                                                                                                                                                                                                                                                                                                                                                          GRVAGAAPPAAAAAAKAAAKAAQYGLGGAGGLGAGGLGAGGLG------AGGLGAG 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      631
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1] SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
MEDLINE-91104868; PubMed-1702999;
Pierce R.A., Deak S.B., Stolle C.A., Boyd C.D.;
"Heterogeneity of rat tropoelastin mRNA revealed by cDNA cloning.";
Blochemistry 29:9677-9683(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 264-533 AND 558-864 FROM N.A., AND ALTERNATIVE SPLICING
                                                                                                          --GAGAAAAAAAAAAGGAGQGGYG---GLGSQGTSGR-----GGLGGQGAGAAAAAA
                                                                                            -------GQGGYGGLGSQGTSGRGCLGGGGAG
                                                                                                                                              AAAAAAAAAAGGA----GLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                        479 AAG----GAG-----OGGYGGLGSQ-GTSGPGGYGPGQQTSGRGGLGGQG--
                                                                                                                                                                                                                                                                                                                                                                                                                ---AGAA-AAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GY - - GGLGSQGTSGRGGLG - GQGAGAAAAAAAAAAGGAGGGCYGGLGSQGTSGRGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 781-864 FROM N.A.
MEDLINE-88330868; PubMed=2971041;
Deak S.B., Pierce Belsky S.A., Riley D.J., Boyd C.D.;
"Rat tropoelastin 1s synthesized from a 3.5-kilobase mRNA.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        632 GAGAAAAAAAAAGGAGGG-----YGG-LGSQGTSGPGGYG 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=92241859; Pubmed=1572637;
Pierce R.A., Alatawi A., Deak S.B., Boyd C.D.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-0cT-1996 (Rel. 34, Created)
01-0cT-1996 (Rel. 34, Last sequence update)
01-0cT-2001 (Rel. 40, Last annotation update)
Elastin precursor (Tropoelastin) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Biol. Chem. 263:13504-13507(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Rattus norvegicus (Rat).
                                                                                           AAAGGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [2]
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ELS_RAT
Q99372;
             245
                                                                                          238
                                                                                                                    365
                                                                                                                                              269
                                                                                                                                                                                                 322
                                                                                                                                                                                                                                                                                                           430
                                                                                                                                                                                                                                                                                                                                                                                                                519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
                                                               셤
                                     ò
                                                                                          ò
                                                                                                                 8
                                                                                                                                      ò
                                                                                                                                                                   g
                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                           ô
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     38;
                                                   Genomics 12:651-658(1992).

-I- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND NUCHAL LIGAMENT, WHICH MUST EXPAND RAFIDLY AND RECOVER COMPLETELY.

-I- SUBDUIT: THE POLYMERIC ELEASTIN CHAINS ARE CROSS-LINKED TOGETHER INTO AN EXTENSIBLE 3D NETWORK.

-I- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.

-I- ALTERNATIVE PRODUCTS: Various isoforms are produced by alternative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       173 LGAGVGAGGKPGKVPGVGLPGVYPGGVLPGTGARFPGVGVLPGVPTGTGVKAKVPGGGGG 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AFSGIPGVGPFGGQQPGVPLGYPIKAPKLPGGYGLPYTNGKLPYGVAGAGGKAGYPTGTG 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         53 GGLGGGALGPGGKPPKPGAGLLGAFGAGPGGLGGAGPGAGLSYASRPGGVLVPGGGAGAA 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAAAAAAGGGGGGGGGGCT-----GSQG-----TSGRGGLGGQGAGAAAAA 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAYKAAAKAGAGLGGIGGVPGGVGVGGVPGAVGVGGVPGAVGGIGGIGGLGGLGVSTGAVVPO 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .112 AAAG-GAG--QGGYGGLGSQGTSGRGGLGGQGA-------GAAAAAAAAAAAAGGA 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGLGGQGAGAAAAAAAAAAG-----GAGQGGYGGLGSQG----TSGRGG--LGGQGAGAA 68
Elements of the rat tropoelastin gene.associated with alternative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             293 VGSQAAVAAAKAAKAKAGAGGGULPGVGGGGIPGGAGAIPGIGGITGAGTPAAAAAKAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GQGGYGGLGSQGTSG-----PGGYG----PGQQTSGRGGLGGQ-----G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AGAAAAAAAAA---GGAGQG----GYGGLGSQGTSGR----GGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGGAGQGGYGGLGSQGTSGR-GGLGGQGAGAAAAAAAAAAGG-AGQGGYGGLGSQGTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGLGGQGAGAAAAAAAAAAAGAGGGCCGCGC--SQGTSGPGGYGPGQQTSGRGGLGG--Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26.5%; Score 939; DB 1; Length 864;
38.3%; Pred. No. 6.3e-36;
tive 32; Mismatches 279; Indels 198;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MISSING (IN CERTAIN ISOFORMS).
MISSING (IN CERTAIN ISOFORMS).
MISSING (IN CERTAIN ISOFORMS).
456894BB09E79FD4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Structural protein; Connective tissue; Repeat; Signal;
                                                                                                                                                                                                                                                                                                                  LYS.
                                                                                                                                                                                                                                                                                  splicing.
-1- PIM: THE CROSSLINKS ARE MADE OF DEAMINATED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY SIMILARITY.
ELASTIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M60647; AAA42269.1; -.
EMBL; M60372; AAA4226.1; -.
EMBL; M86372; AAA42271.1; JOINED.
EMBL; M86363; AAA42271.1; JOINED.
EMBL; M6364; AAA42271.1; JOINED.
EMBL; M6364; AAA42271.1; JOINED.
EMBL; M6364; AAA42271.1; JOINED.
EMBL; M6364; AAA42271.1; JOINED.
EMBL; M6376; AAA42271.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            72786 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; AAA42272.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21
864
307
308
823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               263
308
308
809
864 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M86375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 316;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
   ARAKA MARAKA MAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

Dec

```
SWISS-PROT entry is copyright. It is produced through a collaboration sen the SWISS Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Fleischmann R.D., Alland D., Elsen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                     459
                                                                    422
                                                                                                                                                                                                                                   704
                                                                                                                                                                                                                                                                   571
                                                                                                                                                                                                                                                                                  s:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-HJRW;
MEDLINE-98295807; PubMed-9634230;
MEDLINE-98295807; PubMed-9634230;
Gordon S.V., Elgimeier K., Gas S., Barry C.E. III, Tekaia F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Deviln K., Fellvell T., Gentles S., Handlin N., Holroyd
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
                                                                                                                                                      GVPGAGTPAAAAAAAAAAAAAAAAAAGYGLGPGVGGVPGGVGVGGLPGGVGPGGVTGTGTGP
                                                                                                                                                                                                                     LGGQGAGAA - - AAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGA
     --GSQGTSGRG
                                471 AAAAAAAKAAKYGAGGAGTLGGLVPGAVPGALPGAVPGALPGAVPGAPGTG
                                                                                                                                -----RGGLGGQGAGAAAAAAAA------AAAG-GAGQGGYG------GLGSQGTS
                                                                ----AGGAGQGGY-GGLGSQGTSG----
                                                                                                                                                                                                 GRGGLGGQGAGA----AAAAAAAAAAAGGAGQGGYGGLGSQ-GTSGPGGYGPGQQTSGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             iitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Whole genome comparison of Mycobacterium tuberculosis
                                                                                                                                                                                                                                                                                                                                                                                                GGQGAGAAAAAAAAGGAGQGG----YGG-LGSQGTSGPGGYG 667
                                                                                                                                                                                                                                                                                                                                                                                                                               812 ARPGFGLSPIYPGGGAGGLGVGGKPPKPYGGALGALGYQGGGCFG 856
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv1325c precursor.
RV1325C OR MT1367 OR MTCY130.10C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    603 AA
GAGAAAAAAAAAGGAGQGGYGGL-
                                                                GLGGGGAGAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-CDC 1551 / Oshkosh;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               laboratory strains.";
Submitted (APR-2001) t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YD25_MYCTU
Q10637;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bishai W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWI
between
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
YD25_MYCTU
                                                                                                 531
                                                                                                                                                                591
                                                                                                                                                                                               460
                                                                                                                                                                                                                               651
                                                                                                                                                                                                                                                                                                                               572
                                                                                                                                                                                                                                                                                                                                                               756
                                                                                                                                                                                                                                                                                                                                                                                                629
                                                                387
                                                                                                                                423
                                                                                                                                                                                                                                                                                                705
                              g
                                                                                            g
                                                                                                                                ò
                                                                                                                                                             g
                                                                                                                                                                                               ö
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                             á
                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
```

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce.or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                29;
                                                                                                                                                                                                                                                                                                                                                                                                  196
                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            586
                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                      87
                                                                                                                                                                                                                                                                                                                                                             65
                                                                                                                                                                                                                                                                                                                                     GSTLGAANAAALGP-----YTELLAAGADEVSAAIASLFAAHGQA----YQAV
                                                                                                                                                                                                                                                                                                                                                                                    AAAAAAGGA - - GQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GTSGRGGLGGQGAGAAAAAAAAAAAGGAG----QGGYGGLGSQGTSGRGGL---GGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---GAGGAGGR-SGLFEVLAAGGAGGTGGLSVNGGTGGTG--GTGGGGGLFSNGGAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         291 GFGVSGSAGGNGGTGGDGGIFTGNGGTGGTGGTGGTGTGNQLVGGEGGAGGAGGNAGILFGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GDGGAGGAG-----GTGIEFGSVGGAGGAGGNAAGLSGAGGAG----GGFG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --AGMVGNGGDGGAGGASVVANGGVGSGGN----ATLIGNGGNGGNGGNGGNGGNGGNGGNGG
                                                                                                                                                                                                                                                                                                                                                                                                                                  144 AAAAAAAAGGA---GQGGYGGLGSQG-TSGPGG---YGPGQQTSGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           308 AAAAAAAAAGGAGGGYGGL--GSQGTSGPGGYGPGQQ-TSGRGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          365 AAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLG-SQGTSGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                405 GCIGGSG---GSGGAAKLIGDGGAGGTGGDSVRGAAGSGGTGGTG-----GLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQCAGAAAAAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGYGGLGS-QGTSG
                                                                                                                                                                  POTENTIAL.
HYPOTHETICAL PE-PGRS FAMILY PROTEIN RY1335C.
GLY-RICH.
                                                                                                                                                                                                                                                                                                               Indels 126;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -ETAGDGGAGGN-----AGLLNGDGGAGGAGGLGIAGDGGNGGKGGK----
                                                                                          TIGR; MT1367; -.
Tuberculist; Rv1325c; -.
Tuberculist; Rv1325c; -.
Pfam; PF00934; PE_region.
Probom; PF00934; PE_1.
Probom; P0001223; PE_region; 1.
Hypothetical protein; Repeat; Signal; Complete proteome.
                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                 G -> V (IN REF. 2).
G -> D (IN REF. 2).
G -> D (IN REF. 2).
L -> F (IN REF. 2).
t 4F9BCB82B07AE964 CRC64;
                                                                                                                                                                                                                                                                                                              35; Mismatches 256;
                                                                                                                                                                                                                                                                                                 6.5e-33;
                                                                                                                                                                                                                                                                                         DB 1;
                                                                                                                                                                                                                                                                                       Score 869;
Pred. No. 6
                                                                                  EMBL; AE007010; AAK45630.1; ALT_INIT.
                                                                                                                                                                                                                                                                 Ξ.
                                                                                                                                                                                                                                                                                      24.5%;
38.0%;
                                                                                                                                                                                                                                                                49575
                                                                      EMBL; Z73902; CAA98089.1;
                                                                                                                                                                                                                                                                                                               Matches 256; Conservative
                                                                                                                                                                   30
                                                                                                                                                                                                    603
132
135
337
508
                                                                                                                                                                                                    114
132
135
135
337
508
603 AA;
                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                      CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                    Best Local
                                                                                                                                                                    GNAL
                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                     28
                                                                                                                                                                                                                                                                                                                                                             22
                                                                                                                                                                                                                                                                                                                                                                                                          99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            535
 δ
                                                                                                                                                                                                                                                                                                                                                           D
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ă
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
```

PGGYGPGQQTSGI 675

```
387 GLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAG 446
                                                                                                                                                                QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQ 506
                                                                                                                                                                                                                                507 QTSGRGGLGGQGAGAAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAA 566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TISSUE-Nuchal ligament;
MEDLINE-89274159; Pubmed-2543440;
Yeh Hr; Anderson N., Ornstein-Goldstein N., Bashir M.M.,
Rosenbloom J.C., Abrams W.R., Indik Z., Yoon K., Parks W., Mecham R.,
Rosenbloom J.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Structure of the bovine elastin gene and S1 nuclease analysis of alternative splicing of elastin mRNA in the bovine nuchal isgament."; Biochemistry 28:2365-2370(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Manohar A., Shi W., Anwar R.A.; "Partial characterization of bovine elastin gene; comparison with the
   AG-----GGAG-GGYGGGGEHGGGGGGGGGG-GAGGGYGAGGEHGGGAGGGQGGG 238
                                                                                                                                                                                                ----GGYGAGGEHGGGGGGGGGGGGGAAAVGEHG---GGYG--GGQGGGDGGGYGTGG 357
                                                                                                                                                                                                                                                                                                                              --GDHGAAGYGG------GEGGGGGGGGGGG-----YGDGCAHGGGYGG-----GAG 435
                                                                                                                                 275 AGGCQG------GGAG----GGYCAGGEHGGGAGGCQG-------GGAG
                                  LGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGGGGYGGLGSQGTSGRG
                                                      239 AG------GGYGAGGEHGGAG-GGQG------GGAG-----GGYGAGGEHGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostom1;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
                                                                                                                                                                                                                                                               358 EHGGGYG-GGOGGG---AGGGYGTGGEHGGGYGG-----GOGGGGGYGAG-----
                                                                                                                                                                                                                                                                                               567 AAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGSQGTSGRG
                                                                                                                                                                                                                                                                                                                                                              627 GLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYG 667
                                                                                                                                                                                                                                                                                                                                                                                   b, and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P04385; P04986; P04987; Q29421;
13.AUG-1987 (Rel. 05, Created)
13.AUG-1987 (Rel. 05, Last sequence update)
16.CCT-2001 (Rel. 40, Last annotation update)
Elastin precursor (Tropoelastin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 747 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Raju K., Anwar R.A.;
"Primary structures of bovine elastin a,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sequences of cDNA clones.";
J. Biol. Chem. 262:5755-5762(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1-27 FROM N.A. MEDLINE-91234332; Pubmed-2031719;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-87194772; Pubmed-3032943;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1-27 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               taurus (Bovine).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ELS_BOVIN
                                  327
 191
                                                                                                                                                                                              307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                             g
                                                                                                 ò
                                                                                                                               8
                                                                                                                                                                 ð
                                                                                                                                                                                             셤
                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics. The Three are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           33;
                                                                                                                                                                                                                                                                                                                                                                                                         -i- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
-i- SUBCELLULAR LOCATION: Cell wall (Potential).
-i- DOMAIN: THE REPEATED DOMAINS OF THE PROTEIN FORM A BETA-PLATED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          266
                                                                                                                       01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
01-OCT-1994 (Rel. 30, Last annotation update)
61ycine-rich cell wall structural protein 1.8 precursor (GRP 1.8).
Phaseolus vulgaris (Kidney bean) (French bean).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1; Fabales; Rabaceae; Papilionoideae; Phaseoleae; Phaseolus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88 GSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       G--GYAGEHGVVGYGGGS------GGGQGGVGYGGDQGAGYGGGGGGSGGG----G 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       152 GGGYNAGGAOG-----GGYGTG-----GGAGGGGGGGG----GDHGGGYGGGQG 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY: THE N-TERMINAL SIGNAL SEQUENCE OF ABOUT 30 AA SHOWS ABOUT 60% HOMOLOGY TO THAT OF THE GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN GRP 1.0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19 GVCSARRALLTLDAGYGLGHGTG------GGYGG------AAGSYGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        207 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Pred. No. 2.7e-32;
16; Mismatches 170; Indels 199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 1.8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             _
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 465;
                                                                                                                                                                                                                                                                                                        STRAIN-cv. Tendergreen;
MEDLINE=89991109; PubMed=3208742;
Keller B., Sauer N., Lamb C.J.;
"Glycine-rich cell wall proteins in bean: gene structure and association of the protein with the vascular system.";
EMBO J. 7:3625-3633(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X 22 AA TANDEM REPEATS.
B5C4A9B983B43607 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .7e-32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; X13596; CAA31932.1; -.
PIR; S01820; S01820.
Cell wall; Structural protein; Repeat; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8 X 22 AA TANDEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            =
=
=
-
                                                                                               465 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 853.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLY-RICH.
                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            _
=
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          36683 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24.1%;
39.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 39.9
Matches 256; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                         SHEET CONFIGURATION.
                                                                                               STANDARD:
33 4
205 3
465 AA;
                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                              eurosids I; Faba
NCBI_TaxID-3885;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31
                                                                                               GRP2_PHAVU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                             P10496:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                              GRP2_PHAVU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         267
               셤
                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

c deduced from the

```
EBN1_EBV
                      293
                                                              349
                                                                                                                       406
                                                                                                                                           458
                                                                                                                                                               466
                                                                                                                                                                                                                                                                                                         099
                                                                                                   407
                                                                                                                                                                                    510
                                       q
                      ò
                                                                              g
                                                                                                     å
                                                                                                                       8
                                                                                                                                             ð
                                                                                                                                                             q
                                                                                                                                                                                   à
                                                                                                                                                                                                       g
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                            DD
                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@iBp-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGQGAGAAAAAAAAAG-GAG--QGGYGGLGSQGTSGPGGY--GPGQQTSGRGGLGG--Q 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -GAFPGGF------FGAGGGAAGAAAYKAAAKAGAAGLGVGGI------GGV 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----GAGQGGY-GGLGSQGTSGRGGLGGQGAGAA-AAAAAAGGAG---QGGYGGLGSQ 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20 GGLGGQGAGAAAAAAAAAAGAGGGGYGGLG---SQGTSGRGGLGGQGAGAAAAAAAA 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGVPGAVPGGVPGGVFFPGAGLGGLGVGGLGPGVKPAKPGVGGLVGPGLGAEGSALP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GAGVKPKAQVGAGAFAGIPGVGPFGGQQP-----GLPLGYPIKAPKLPAGYGLPYKTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27; Mismatches 295; Indels 124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 777.5; DB 1; Length 747;
Pred. No. 8.2e-29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       MISSING (IN ELASTIN B).
MISSING (IN ELASTIN C).
MRS - MAG (IN REF. 2 AND 3).
E -> G (IN REF. 2 AND 3).
E 533C03E411643DB3 CRC64;
                                                                                                                                                                      InterPro; IPR003979; tropoelastin.
PRINTS; PR01500; TROPOELASTIN.
Structural protein; Connective tissue; Repeat; Signal;
                                                                                                                                                                                                                       ELASTIN.

OXIDATIVE DEAMINATION.

OXIDATIVE DEAMINATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      =
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     = = - = =
                                                                                      EMBL; JO2717; AAA30503.1; --
EMBL; K03505; AAA30505.1; --
EMBL; K03506; AAA30705.1; --
EMBL; J02055; AAA30776.1; --
EMBL; M58652; AAA03519.2; --
PIR; B26728; B26728.
PIR; B26728; B26728.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64229 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              39.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                         272
324
400
404
704
                                                                                                                                                                                                    Alternative splicing.
SIGNAL 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               747 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                 645
649
685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                MOD_RES
MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                            MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                        MOD_RES
MOD_RES
MOD_RES
MOD_RES
MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                            MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                              MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                         MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                   MOD_RES
                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              124
                                                                                                                                                                                                                                                                                    ΨÔD,
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
```

```
MEDLINE-84270667; PubMed-6087149;
Baer R., Bankler A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
Glisson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
Tuffnell P.S., Barrell B.G.;
"DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
Nature 310:207-211(1984).
236 KLPYGFGPGGVAGSAGKAGYPTGTGVGPQAAAAAAKAAAKLGAGGAGVLPGVGVGGGGPGIP 295
                                                                                                                                                                                                                                                                   356 GVGVPGVGVP-----GVGVPGVGVPGVGVPGV-GVPGVGVPGALSPAATAKAAAKA 405
                                                                                                                                                                                                                                                                                                                                              457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-90266473; Pubmed-2161150;
Petti L., Sample C., Kieff E.;
Subunciear localization and phosphorylation of Epstein-Barr virus
latent infection nuclear proteins.";
Virology 176:563-574(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 1-26 FROM N.A.
MEDLINE-86259739; PubMed-3460083;
MEDLINE-86259739; PubMed-3460083;
Sample J., Hummel M., Braun D., Birkenbach M., Kleff E.;
"Nucleotide sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: a probable transcriptional initiation site.";
Proc. Natl. Acad. Sci. U.S.A. 83:5096-5100(1986).
                                                                   GTSGR----GCLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRG
                                                                                                                                   GAPGAIPGIGGIAGVGAPDAAAAAAAAAKRAKFGAAGGLPGVGVPGVGVPGVGVPGVGVP
                                                                                                                                                                                                       GLGGQGAGAAAAAAAAAAAGGAGQGGYG--GLGSQGTSGRGGLGGQGAGAAAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                           AKFGARGAVGIGGIPTFGLGPGGFPGIGDAAAAPAAAAKAAKIGAGGVGALGGVVPGAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GIGLGPGGVIGAGVPAAKSAAKAAKAQFRAAAGLPAGVPGLGVGAGVPGLGVGAGVPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LGVGAGVPGPGAVPGTLAAAKAAKFGPGGVGALGGVGDLGGAGIPG-GVAGVVPAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 604 AAAGGAGQGGYGGLGSQGTSGRGGLGGQ----GAGAAAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --TSGRGGLGG-QGAGAAAAAAAAAGGAGQGGYG-GLGSQGTSG--PG-GYGPGQQTS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----GRGGLGGQGAGAAAAAA------AAAG------GAGQGGYG-GLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             546 TSGRGGLGGQGA--GAAAAAAAAAGGAGQGYGGLGSQGTSGRGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                           AGQGGYG--GLGSQGTSGRGCLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Epstein-barr virus (strain B95-8) (Human herpesvirus 4).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Gammaherpesvirinae; Lymphocryptovirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 469-607
MEDLINE-96006523; PubMed-7553871;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            641 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EBNA-1 nuclear protein. BKRF1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    705 PIGGGAGGLGVG 716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    T--SGPGGYGPG 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBCELLULAR LOCATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=10377;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63 GPRHRDGVRRPQKRPSCIGCKGTHGGTGAG--AGAGGAGGAGAGGAGGAGGAGGAGGAG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       223 GLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  283 QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQ 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          343 QTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        175 GPGQQTSGR------GGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRG 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAA 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OSSSSGSPPRRPPPGRRPFFHPVGEADYFEYHOEGGPDGEPDVPPGAIEQGPADDPGEGP 446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GAGQGGYGGLGSQG-TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGY 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                               SOMEWHAT
                                                                                                                                                 FUNCTION: INVOLVED IN LATENT CYCLE. EBNA-1 FUNCTIONS IN THE MAINTENANCE REPLICATION OF EBV EPISOME. TRANSACTIVATING FACTOR FOR THE ORIGIN AND ENHANCER FUNCTIONS OF ORIP.
SUBCELLULAR LOCATION: UNCLEAR. FREE IN THE NUCLEOPLASM, SOMEWHIS ASSOCIATED WITH THE CHROMATIN AND HARDLY, IF AT ALL ASSOCIATED WITH THE NUCLEAR MATRIX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nuclear protein; DNA-binding; Transcription regulation; Activator;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Edwards A.M., Frappler L.;

"Crystal structure of the DNA-binding domain of the Epstein-Barr

"Crystal structure of the BNA-binding domain of the Epstein-Barr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21.5%; Score 763; DB 1; Length 641;
35.7%; Pred. No. 3.3e-28;
tive 22; Mismatches 209; Indels 186;
A., Barwell J.A., Pfuetzner R.A., Furey W.F. Jr.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       352 GLY/ALA-RICH.
56427 MW; 4D161653E16FC341 CRC64;
                                                                                         virus origin-binding protein EBNA 1.
Cell 83:39-46(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M13941; AAA45889.1; -. PIR; A03773; QQBE31. PIR; S33021; S33021.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; V01555; CAA24816.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF02905; EBNAl; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PDB; 1VHI; 23-DEC-96.
TRANSFAC; T00211; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     641 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            232;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             387
   RAHARA RA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING UNIT (G-X Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
                                   STGPRGQGDGGRRKKGGWFGKHRGQGGSNPKFENIAEGLRALLARSHVERTTDEGTWVAG 506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kimura S.;
"Partial characterization of cDNA clones encoding the three distinct pro alpha chains of type I collagen from rainbow trout.";
Fisheries Sci. 64:780-786(1998).
-i- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (FIBRILLAR FORMING COLLAGEN).
-!- SUBUNIT: TRIMERS OF ONE ALPHA 2(1) AND TWO ALPHA 1(1) CHAINS.
-!- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM HYDROXYAPATITE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AMINO-TERMINAL PROPEPTIDE (POTENTIAL)
COLLAGEN ALPHA 2(1) CHAIN.
CARBOXYL-TERMINAL PROPEPTIDE
614 YGGLGSQGTSGRGGLGG----QGAGAAAAAAAAGG------AG
                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Verrebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC. . .) (POTENTIAL).
W; 7BBZF1F80DB10C93 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Saito M., Kunisaki N., Hirono I., Aoki T., Ishida M., Urano N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Repeat; Hydroxylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Salto M., Takenouchi Y., Kunisaki N., Kimura S.;
"Complete primary structure of rainbow trout type I collagen
consisting of alphal(I)alpha2(I)alpha3(I) heterotrimers.";
Eur. J. Biochem. 268:2817-2827(2001).
                                                                          --GPGGY--GPGQQTSGIR
                                                                                               ||| : ||| || 507 VEVYGGSKTSLYNLRRGTALAIPQCRLTPLSRLPFGMAPGPGPQPGPLR
                                                                                                                                                                                                                                                                                                                                       Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
                                                                                                                                                                                                                                                           16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Collagen alpha 2(I) chain precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BY SIMILARITY)
                                                                                                                                                                                                          ٤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Probom; P0002078; Fib_collagen_C; 1.
SMART; SM00038; COLFI; 1.
Extracellular matrix; Connective tissue;
                                                                                                                                                                                                      PRT; 1356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AB052837; BAB55663.1; -.
EMBL, AB008372; BAA33379.1; -.
InterPro; IPR000087; Collagen.
InterPro; IPR000885; Fib_collagen_C.
                                                                                                                                                                                                                                                                                                                                                                                                                               [1]
SEQUENCE FROM N.A.
MEDLINE-21257802; PubMed-11358497;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           126985 MW:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Glycoprotein; Collagen; Signal
                                                                                                                                                                                                                                         16-OCT-2001 (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 417-1356 FROM N.A.
                                                                      649 OGGYGGLGS-----QGTS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF01391; Collagen;
Pfam; PF01410; COLFI; 1.
                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ΑĄ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE-Fibroblast,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1097
                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-8022
                                                                                                                                                                                                    CA21_ONCMY
093484:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROPEP
                                                                                                                                                                                   CA21_ONCMY
                                                                                                                                                                 RESULT 15
                                   g
                                                                    ð
                                                                                                         a
```

```
32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95 RGGLGGQGAGAAAAAAAAAAGGAGQGG-YGGLGSQGTSGRGGLGGQGAGAAAAAAAAG 153
                                                                                                                                                                                                                                                                                          275 GRGEPGINGAVGPVGPVGNPGNNGINGAKGAAGLPGVAGAPGFPGPRGCPGPGPGPGGPGGSTC 334
                                                                                                                                                                                   154 GAGQGGYGGL------GSQGTSGPGGY--GPGQQ-------TSGRGG 185
                                                                                                                                                                                                                                   186 LGGQGAGAAAAAAAAAGGAG----QGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAG 241
                                                                                                                                                                                                                                                                                                                                                           574
                                                                                                                                                                                                                                                                                                                                                                                 L-----GSQGTSG-RGGLGGQGAGAAAAAAA--GGAGQ-GGYGGLGSQGTSGPGG 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGG-----YGGLGSQGTSGRGGLG 427
                                                                         GAGQGGYGGL------GSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG 288
                                                                                                                                                                                                                                                                                                                                                                                                             634
                                                                                                                                                                                                                                                                                                                                                                                                                                       694 GPAGPAGLRGPSGERGEGGPAGLPGFAGPPGSDGOSGPRGEKGPAGGKGDVGPAGPAGPS 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GOGAGAAAAAAAAAAGGAGGGYGGL-GSQCTSGRGGLGGQGAGAAAAAAAAAAAGGAGQG 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GYGGL------GSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAGGAGQG 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GPRGLRGDPGPGGPQGEQGVVGPAG-----ISGDKGPSGESGPPGAPGTAGPQGVLGPS 864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        537 GYGGL----GSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAGGAGQGGYGGL-GSQGTSGRGGLG 591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         592 GQGAGAAAAAAAAAAGGAGQ-GGYGGLGSQGTSG-------RGGLG--GQGAGA 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  636 AAAAAAAAAGGAG----QGGYGGLGSQGTSG-----PGGYGPGQQTSGIRIRAPS 681
                              Gaps
                                                     9 QMGRGSMASGRGGLGGQGAGAAAAAA--AAAGGAG-----QGGYGGLGSQGTSG-RG 58
                                                                                                          GLGGQGAGAA------AAAAAAAAAAAAGGAG-----QGGYGGLGSQGTSG 94
                                                                                                                                                                                                                                                                                                                                              GNDGPPGRPGAAGFKGDRGEPGSPGALGSSGQPGPNGPAGSAGRPGNRGESGPTGNGGPV
                            353; Indels 177;
     DB 1; Length 1356;
  21.1%; Score 747.5; DB 30.6%; Pred. No. 2.4e-27; tive 53; Mismatches 355
Query Match
Best Local Similarity 30.6
Matches 257; Conservative
                                                                                                            59
                                                                                                                                                                                                                                                                                                                                                        515
                                                                                                                                                                                                                                                                                                                                                                                   289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       925
                                                                                                                                                                                     q
                                                                                                            å
                                                                                                                                    g
                                                                                                                                                              ò
                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
```

Search completed: December 18, 2002, 16:18:15 Job time : 21.6063 secs

```
GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2002, 16:13:04; Search time 40.5691 Seconds

(without alignments)

3509.535 Million cell updates/sec
```

Title: US-09-490-291-6
Perfect score: 3543
Sequence: 1 MASMTGCQQMGRGSMASGRG......TSGIRIRAPSTSFEHHHHHH 691
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : SPTREMBL_21:*

1: Sp_archea:*
2: sp_bacteria:*
3: sp_lungi:*
4: sp_human:*
5: sp_nmammal:*
7: sp_nmammal:*
7: sp_nmammal:*
9: sp_nmammal:*
10: sp_nde:*
10: sp_nde:*
11: sp_rodent:*
12: sp_rius:*
13: sp_virus:*
14: sp_wirus:*
15: sp_wirus:*
16: sp_mammal:*
16: sp_mammal:*
17: sp_rodent:*
18: sp_rodent:*
18: sp_virus:*
19: sp_virus:*
16: sp_virus:*
16: sp_virus:*
16: sp_virus:*
16: sp_virus:*
16: sp_virus:*
17: sp_virus:*
18: sp_virus:*
18: sp_virus:*
19: sp_virus:*
10: sp_virus:*
10: sp_virus:*
10: sp_virus:*
10: sp_virus:*
10: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

sp_archeap:*

Q8wsw4 nephila cla O46172 nephila cla O46171 nephila cla Q9biu3 doLomedes tr Q9biu4 dolomedes tr Q9bit7 argiope tri Q9bit7 nephila mad O17434 nephila cla O766 aftheraea p Q964f4 antheraea y Q8wiy9 mycobacteri mycobacteri mycobacteri mycobacteri mycobacteri mycobacteri Description 08v120 1 053559 1 053577 0 053775 1 SUMMARIES 046172 046171 09BIU3 09BIU7 09BIU7 017434 076786 0964F4 08VIX9 053559 053557 053775 08VIZ1 ВП Query Match Length 1715 1489 1079 1306 1729.5 1594.5 1568 1539.5 1452.5 1450 1386 1348.5 1273 1260.5 1242 1198.5 1198 Score 112110987654 Result

Q9bivl argiope aur O53552 mycobacteri O53215 mycobacteri Q9biu8 argiope tri Q9hhw2 nephila mad	Q8vj23 mycobacteri O53395 mycobacteri O02402 pinctada fu Q8vki7 mycobacteri	Oggubs galleria me Ob6810 mycobacteri O53439 mycobacteri O53844 mycobacteri O44359 pophila ola	Q8vkd2 mycobacteri O44358 ncphila cla O53845 mycobacteri O9bit5 nephila mad O53458 mycobacteri O53484 mycobacteri	O99109 argiope tri O53809 mycobacteri O16987 araneus dia O80115 mycobacteri O53156 mycobacteri O53556 mycobacteri O6808 mycobacteri
Q9BIV1 O53552 O53215 Q9BIU8 Q9NHW2	08VJ23 053395 002402 08VK17	09GUB5 006810 053439 053844 044359	Q8VKD2 Q44358 Q53845 Q9BIT5 Q53435	Q9BIU9 Q169BT Q169B7 Q8VJ15 O50415 O53556 P949B6
5 16 5 5	16 16 16 7	10 10 10 10	16 16 16	16 16 16 16 16
447 1381 1665 1002 1884	1507 1538 738 1408	1468 1329 853 749	871 871 882 863 767	651 783 636 775 731 714 741
33333 3233 9003 9003			200.6 200.6 200.6 200.6	22882 22882 22886 272 200 200 200 200 200 200 200 200 200
1197 1196.5 1180 1167	1158.5 1158.5 1157 1157	1152.5 1152 1151 1114.5	1088.5 1084 1083 1045 1032.5	1015.5 1014 1013 1008 1007 992 989
17 18 19 20 21	22 23 24 25	27 28 29 30 31		8000178444 8000178444

SSU	RESULT 1					
S	Q8WSW4					
	Q8WSW4 PRELI	PRELIMINARY;	PRT;	644 AA.	Α.	
AC	QBWSW4;					
DŢ	01-MAR-2002 (TrEM	BLrel, 20	, Created)			
ρŢ	01-MAR-2002 (TrEM	BLrel. 20	, Last sed	neuce	update)	
DŢ	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	BLrel. 21	, Last ann	otatio	n update)	
DE	Dragline silk protein (Fragment).	tein (Fra	gment).		•	
SO	Nephila clavipes (Orb spider).	(Orb spid	er).			
႘	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;	a; Arthro	poda; chel	icerat	a; Arachnida;	Araneae;
	Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae;	telegynae	Araneold	ea; Te	tragnathidae;	Nephila.
ŏ	NCBI_TaxID=6915;	1				•
RN	[1]					
RP	SEQUENCE FROM N.A.					
RA	Basel R.;					
RL	Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases	01) to th	e EMBL/Gen	Bank/D	DBJ databases	
DR	EMBL; AY061814; AAL32375.1;	AL32375.1				
DR	InterPro; IPR000104; Antifreeze_1.	04; Antif	reeze_1.			
DR	PRINTS; PR00308; ANTIFREEZEI.	ANTIFREEZ	EI.			
	NON_TER 644	644				
ğ	SEQUENCE 644 AA	.; 52104	MW; 9024A	F2F65C	644 AA; 52104 MW; 9024AF2F65C5F75A CRC64;	

ALIGNMENTS

FT NON_TER 644 644
SQ SEQUENCE 644 AA; 52104 WW; 9024AF2F65C5F75A CRC64;
Query Match 56.8%; Score 2012; DB 5; Length 644;
Best Local Similarity 67.9%; Pred. No. 9e-102;
Matches 465; Conservative 19; Mismatches 63; Indels 138; Gaps

31;

ò

```
g
                                                                                              셤
                                                                                                                                           셤
                                                                                                                                                               à
                                                                                                                                                                                   q
                                                                                                                                                                                                         à
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                  ò
                                                                           ò
                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SECUENCE FROM N.A.
MEDLINE=94165058; PubMed=8120021;
Beckwitt R., Arcidiacono S.;
"Sequence conservation in the C-terminal region of spider silk proteins (Spidroin) from Nephila clavipes (Tetragnathidae) and Araneus bicentenarius (Araneidae).";
J. Biol. Chem. 269:6661-6663(1994).
                                                                                                                       419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-98318805;
MEDLINE-98318805;
MEDLINE-98318805;
MEVOLUTE R., Arcidiacono S., Stote R.;
Evolution of repetitive proteins: spider silks from Nephila clavipes (Tetragnathidae) and Araneus bicentenarius (Araneidae).";
Insect Blochem. Mol. Biol. 28:121-130(1998).
                                                                                                                                                              TSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAA 479
                                                                                                                                                                                   GCQCAGAAAAAAAAAAGGACQCGYCGLGSQCTSGPGGYGP-GQQTSGRCGLGGQGAGAAA 359
                                                                                                 301
                                                                                                                                  Gaps
                                           AGGAGOGGYGGLGSQGTSGPGGYGP-GQQTSGRGGLGGQGAGAAAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                 GGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAA
                                                                                                                                                                                                                                                                                           GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL
                                                                                                                   AAAAAAAAGGGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Araneae;
Nephila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 156;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota, Metazoa, Arthropoda, Chellcerata, Arachnida,
Araneomorphae, Entelegynae, Araneoidea, Tetragnathidae,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    617 AA; 49410 MW; 1EF0CE9269A832E2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                       046172;
01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-JUN-2001 (TrEMBLrel. 19, Last annotation update)
Dragline silk protein spidroin i (Fragment).
Nephila clavipes (Orb spider).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        56.4%; Score 1998; DB 5; 1
66.6%; Pred. No. 5e-101;
tive 18; Mismatches 58;
                                                                                                                                                                                                                                                                                                                                                                                                                Ā
                                                                                                                                                                                                                                                                                                                                                                                                                617
                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                   GTSGPGGYGPGQQTSGIRIRAPSTS 683
AGREGLEGQCAG---AAAAAAAGGAGQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 462;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                046172
                                                                                                                                                              420
                                                    195
                                                                         301
                                                                                              249
                                                                                                                    360
                                                                                                                                         302
                                                                                                                                                                                  348
                                                                                                                                                                                                        480
                                                                                                                                                                                                                            393
                                                                                                                                                                                                                                                 539
                                                                                                                                                                                                                                                                      445
                                                                                                                                                                                                                                                                                                                                                                                          RESULT
046172
                                                                                                                                                                                g
                                                                                                                                                                                                                         g
           g
                                                   g
                                                                                             g
                                                                                                                                      g
                                                                                                                                                             ð
                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                         à
                                                                                                                   à
                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                          ò
```

ASMTGGOOMGRGSMASGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLG 61

ò

```
GOGAG---AAAAAAAAGGVGQ------GGLGGGGGGGGGGAGAAAAAAAGGAGGGG 100
                                                                                                     181
                                                                                                                                                            231
                                                                                                                                                                                       200
                                                                                                                                                                                                                   291
                                                                                                                                                                                                                                              244
                                                                                                                                                                                                                                                                          350
                                                                                                                                                                                                                                                                                                     292
                                                                                                                                                                                                                                                                                                                                 410
                                                                                                                                                                                                                                                                                                                                                            335
                                                                                                                                                                                                                                                                                                                                                                                         470
                                                                                                                                                                                                                                                                                                                                                                                                                   384
                                                                                                                                                                                                                                                                                                                                                                                                                                                 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                57
122 YGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTS
                                                                                                                 101 YGGLGSQG-AGRGGSGGGAGG----AAAAAAGGAGQGGYGGLGSQG------A
                                                                                                                                                                         GGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAAAGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                      590 LGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQ
                                         GOGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                         QGTSGRGGLGGQGAGAAAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGP-GQQTSGRGGL
                                                                                                                                                                                                                                                                                         GYGGLGSQCTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                ----AAAAAAGGAGQGGYGGLGGQG-AGQGGYGGLGSQGSGRGGLGGQGAG----AAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       530 AGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGRGG
                                                                                                                                                         182 GRGGLGGQGAGAAAAAAAAAAG------GAGQGGYGGLGSQGTSGRGGLGGQGAGA
                                                                                                                                                                                                                  232 AAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Araneae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEGUENCE OF 449-544 FROM N.A.
MEDILE-991148687; Pubmed-9487707;
Arcidiacono S., Mallo C., Kaplan D., Cheley S., Bayley H.;
"Purification and characterization of recombinant spider silk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "structure of a protein superfiber: spider dragline silk.";
Proc. Natl. Acad. Sci. U.S.A. 87:7120-7124(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JUN-1998 (TEMBLEEL. 06, Last sequence update)
01-DEC-2001 (TEMBLEEL. 19, Last annotation update)
Spidroin 1 (Fragment).
Nephila clavipes (Orb spider).
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; NCBL_TaxID=6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appl. Microbiol. Biotechnol. 49:31-38(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGYGGLGSQGTSGPGGYGPGQQTSGIRIRAPSTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDILTRE-9034959; PubMed-2402494;
XX M., Lewis R.V.;
"Structure of a protein superfiber:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            expressed in Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     046171;
01-JUN-1998
              m
                                                                                                                                                                                                                                                                         292
                                                                                                                                                                                     143
                                                                                                                                                                                                                                                                                                                                 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     650
                                         62
                                                                      28
                                                                                                                                                                                                                                            201
                                                                                                                                                                                                                                                                                                                                                                                                                                              471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 3
046171
```

```
435 AGGGGFGPYGSSYQSSTSYSVTSQGAA--GGLGGYGQGSGAGAAAGAAAGAAGGGGGGGYGQG 492
                                                                                                                                                                                                                                                                                                                      GQGAGAAAAAAAAAAGGAGQGGYGGLG********SQGTSGRGGLGGQGAGAAAA 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLG-GQGAGAAAAAAAAAG-GAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241 G-GAGOGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAG-GAGOGGYGGLGSQGTSGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GSQGTSGRGGLG-GQGAGAAAAAAAAAG-GAGQGGYGGLGSQGTSGRGGLG-GQGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               184 GGLGGQGAGAAAAAAAAAG-GAGQGGYGGLGSQGTSGRGCLG--GQGAGAAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPQQTSGRGGLGGQGAGAAAAAAAAAAGG
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
01-bolomedes tenebrosus.
Dolomedes tenebrosus.
Buraryotas, Metacoas, Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Lycosoidea; Pisauridae; Dolomedes.
                                                                                                      SEQUENCE FROM N.A.
MEDLINE=21179804; PubMed=11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Extreme Diversity, Conservation, and Convergence of Spider Silk
Fibroin Sequences.";
Science 291:2603-2605(2001).
                                                                                                                                                                                                                                                                       45.0%; Score 1594.5; DB 5; Length 691; 56.6%; Pred. No. 3.2e-79; ive 39; Mismatches 89; Indels 183;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ..
<u>--</u>
::
                                                                                                                                                                                                                                                691 AA; 56176 MW; CB46EF9F1991F33E CRC64;
                                                                                                                                                                           EMBL; AF350270; AAK30599.1; -.
InterPro; IPR000104; Antifreeze_1.
InterPro; IPR002952; Eggshell.
PRINTS; PR00308; ANTIFREEZEI.
PRINTS; PR01228; EGGSHELL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 56.6%
Matches 406; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACOCCYCCLG-
                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                       24
                                                                                                                                                                                                                                                                                                                                             ~
                                                                                                                                                                                                                                                                                                                                                                    72
                                                                                                                                                                                                                                                                                                                                                                                        62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         632
   g
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                              28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    496
                                Beckwitt R., Arcidiacono S., Stote R.; "Evolution of repetitive proteins: spider silks from Nephila clavipes (Tetragnathidae) and Araneus bicentenarius (Araneidae)."; Insect Blochem. Mol. Biol. 28:121-130(1998).
                                                                                                                                                                                                                                                                                                                                           280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA 316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGCQGAGAAAAAAAAAAAAAGGAGQGGYGG 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGAGQ-----AAAAAAGGAGQGYGG 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGGGGGCGGLGSQGTSGRGGLGGQGAGAAA 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGIR 676
                                                                                                                                                                                                                                                                                 GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGL 136
                                                                                                                                                                                                                                                                                                                               137 GGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAA 196
                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                     GRGGLGGQGAG-AAAAAAAAAAAAGGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA 76
                                                                                                                                                                                                                                                          GYGGIGGOGAGAGAAAAAAGGAGOGGYGGIGSOG----AGRGGOGAG----AAAAAA 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAAAAAGGAGGGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAAAAGGAGGGGYGGLGSQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----QGAGQGGYGGLGSQG-AGRGGLGGQGAG--AAAAAAAAGGAGQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG
                                                                                                                                                                                                              213;
                                                                                                                                                                                   Query Match
48.8%; Score 1729.5; DB 5; Length 544;
Best Local Similarity 61.3%; Pred. No. 1.4e-86;
Matches 409; Conservative 12; Mismatches 33; Indels 213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         251 LGSQG----AGRGGQGAG----AAAAAAGGAGQGGYGG-------
                                                                                                      Beckwitt R.;
Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases
EMBL; U20329; AAC38957.1; -.
HSSP; P10969; IMGT. 1
                                                                                                                                                                 CC611B3551945615 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                SEQUENCE 544 AA; 44107 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                            SEQUENCE FROM N.A.
                        MEDLINE-98318805;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : :| |
LSSPQAS 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        677 IRAPSTS 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9BIU3;
Q9BIU3;
                                                                                                                                                                                                                                    18
                                                                                                                                                                                                                                                                                 11
                                                                                                                                                                                                                                                                                                         53
                                                                                                                                                                                                                                                                                                                                                      93
                                                                                                                                                                                                                                                                                                                                                                            197
                                                                                                                                                                                                                                                                                                                                                                                                   136
                                                                                                                                                                                                                                                                                                                                                                                                                           257
                                                                                                                                                                                                                                                                                                                                                                                                                                                190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
Q9BIU3
ID Q9
AC Q9
  g
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

157

298 253 356 295 415 348 472 398 532 434 577 544

41;

Gaps

윱

```
85 GAGAGAAAATAAGGAGSGQGDYGDQGGLGGYGQGSGAGSATAPAAGGSGFGQGGFGNRGG 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LG--GQGAGA---AAAAAAAAAGGAGGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAA 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             145 KGAYGQSAGAGVGAAATAAAGGAGSGQGGYG-----DQGGLGGYGQAGAGAASAAA 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGAG---QGGY---GGLGSQGT------SGPGGYGPGQQTSGRGGLGGQG 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  249 -GGLGSQGTSGRGGLG--GQGAGAAAAAAAG-GAGQGGYGGLGSQGTSGRGGLGGGG 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -GLGSQGTSGRGGLG--GQGAGAAAAAAAG-GAGQGGY--GGLGSQGTSGRGGLG-- 427
    251 QGAGAGAASAAAGGGDGYGQGYYGDQG--GRGGY-GQGSGAGSATAAAAGGAGFGQGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7 GQQMGRGSMASGRGGLGGQGAGAAAAAAAAAAG-GAGQGGYGGLGSQGTSGRGGLG--GQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64 GAGAAAAAAAAAG-GAGQGGYG--------GLGSQGTSGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAG-GAGQGGY-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                308 QGGYGNQ-----GGLGSFGQGAGAGAAAASAGGAGSGRGGYG------DQGGLGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGAAAAAAAAAAAGGA---GQGGY---GGLGSQGT------SGPGGYGPGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Lycosoidea; Pisauridae; Dolomedes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 43.5%; Score 1539.5; DB 5; Length 854; Best Local Similarity 50.7%; Pred. No. 3.5e-76; Matches 412; Conservative 46; Mismatches 145; Indels 209;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    854
71881 MW; 5C52F1B0AE31C6C2 CRC64;
                                                                                                                                                                                                                                                                                                                             (1)
MEDLINE-21179804; PubMed-11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.,
"Extreme Diversity, Conservation, and Convergence of
                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QTSGRGGLG--GQGAGAAAAAAAAG-GAGQGGYG------
                                                                                                                                                                 ¥
                                                                                                                                                                854
                                                                                                                                                                                                                                                                                                                                                                                                    Fibroin Sequences.";
Science 291:2603-2605(2001).
EMBL, AF350269, AAK30598.1;
InterPro; IPRO00217; Tubuiin.
PROSITE; PS00227; TUBULIN; UNKNOWN_1.
                                                                                                                                                                                            01-JUN-2001 (TrEMBLrel. 17, 01-JUN-2001 (TrEMBLrel. 17, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                          Fibroin 1 (Fragment).
Dolomedes tenebrosus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    854 AA;
                                                                                                                                                                                                                                                                                                           NCBI_TaxID=156846;
                                                   666 YGP 668
                                                                                574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
NON_TER
SEQUENCE
                                                                          572 GGP
                                                                                                                                                                                  Q9BIU4;
                                                                                                                                                                Q9BIU4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              415
                                                                                                                                RESULT 6
Q9BIU4
                                                   ò
                                                                                  g
                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GCYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGY--GGLGS 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QGTSGRGGLGGQGAGAAAAAAAAAGG-AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAA 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   47 --GSAQRGGLG------AGGAGQ-GYG-----AGSGGQGGAGQGGAAAAT 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GAGAGASAAASASGAGSGQGGYG------GQGGYGQGTGAGAASSAGVAVTVGNT 593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAGAGGGG--GYGGLGSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        149 AAAAGG-AGQGGYGGLGSQGTSGPGQYTSGRGGLGGQGAGAAAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGY - - GGLGSQGTSGRGGLGGQGAGAAAAAAAAAGG-AGQGGYGGLGSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QGAGAAAAAAAAAAGGAGGGTGGLGS-QG-TSGRGGLGGQGAGAAAAAAAAAAGG-AGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                411 -SGOGGY------GOGGAAAAAAAAAAASGAGGAGGAGGAGGAGGGAGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           613 GYGGLGSQGTSGRGGLGG-----QGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAAAGGAGQGGY - -GGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGG-AGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GQCAGAAAAAAAAAGG-AGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                 MEDLINE-21179804; PubMed-11283372; MEDLINE-21179804; PubMed-11283372; Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.; Extreme Diversity, Conservation, and Convergence of Spider Silk Fibroin Sequences "; Science 291:2603-2605(2001).
EMBL; AF350266; AAK30595.1; -.
                                                                                                                                                                                                      Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Araneidae; Argiope.
NCBL_TaxID=156845;
                                                                                                                                                                                                                                                                                                                                                                                                                                               1 44.3%; Score 1568; DB 5; Length 648; Similarity 58.5%; Pred. No. 8.1e-78; 18. Conservative 35; Mismatches 128; Indels 112;
                                                                                                                                                                                                                                                                                                                                                                                                                 648 AA; 52887 MW; ACDB67067B2302CB CRC64;
                                                                                                           (TrEMBLrel. 17, Created)
(TrEMBLrel. 17, Last sequence update)
(TrEMBLrel. 19, Last annotation update)
                                                                               648
                                                                                                                                                           Major ampullate spidroin 1 (Fragment)
                                                                               PRT;
                                                                               PRELIMINARY;
                                                                                                                                                                                          Argiope trifasciata
                                                                                              Q9BIU7;
01-JUN-2001
                                                                                                                          01-JUN-2001
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Sim
Matches 388;
                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
                                                                               Q9BIU7
545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          554
                                               Ŋ
                                              RESULT
09BIU7
```

g ò a ò g

ò

ò g ö g à

ð g g

g

ö 셤 ð

à

q

ð

47;

Gaps

84 97 307

```
1261 AAAGSGPGGYGPGQQGPGGPGAAAAAAAAGGGGGGP-GQQGPGGPGAAAAAGPGGYG 1319
                                                                                                                                                                                                                                                          1320 PGQQGPGAAAAAAGSGPGGYGPGQQGPGGPGAAAAAAGRGPGGYGQGQQGPGGPGAAA 1379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1440 QGPGGPGAAAAAAGPGGYGPGAAAAASAGKGPGGYGPGGGGGGGGGGGGAAAAAGKG 1499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1618 GGYGPGQQGPGGPGAAAAAAAGRGPGGYGPGQQGPGGSGAAAAASGRGPGGYGPGQQGPG 1677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1856 QQGTGAAAAAAGSGAGGYG-----PGQQGPGGPGAAAAAGPGGYGPGQQGPGAAAA 1908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               493 SQG----TSGPGGYGPGQQ-----TSGRG------GLGGQGAGAAAA---- 525
                                                                                                                                                                                                                                                                                                                                                               1380 AAAGPGGYGPGQQGPGAAAAAAAAGSGPGGYGPGQGPGRSGAAAAAAAAAGRGPGGYGPGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1797 AAAGRGPGGYGP-GQQGPGGSGAAAAAGRGPGGYGPGQQGPGGPGAAAAAAGPGGYGPG
                                                                                                                                                                                                          LGGQGAGAAAA-----GLGSQG----GLGSQG----GLGSQG---
                                                                                                                                                                                                                                                                                                                                                                                                               -----GLGGQGAGAAAAA---------AAAAAGGAGQGGYG----GLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     526 ------AAAAAGGAGQGGYGGLGSQGTSGRG-----GLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -------AAAAAAGGAGGKGGLGSQGTSGRG-------G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAGGAGQGGYGGLGSQGTSGRG-------GLGGQGAGAAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----AAAAAAGGAGGGGGGGGGGGGGGGGGGGGAGAA-----AAAA
                                                                                                       ----SGRGGL--GGQGAGAAAAAAAAAGGAGGGYGGLGSQGTSGRG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             643 AAGGAGOGGYG----GLGSQGTS-----GPGGYGPGQQTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ź
       -----AAAAAGGAGQGGYG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1909
                                                                                                       257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             601
                                                                                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  970 PGAAAAAAGSGPGGYGP-GQQGPGCPGAAAAAGRGPGGYGPGQQGPGGPGAAAAA- 1027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1028 GRGPGGYGPGQQGPGGSGAAAAAGRGPGGYGPGQQGPGGPGAAAAAGPGGYGP 1087
                                                                                                                                                                                                                                                                                                             645
                                                                                                                                                                                                                                                                                                                                       145
                                                                                                                                                                                                                                                          691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             71 --AAAAAAGGAGGGYGGLGSQGTSGRG------GLGGGGGAGAAAAAAAA 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----AAAAAAGGAGQGGYG----GLGSQGTS-----GPGGYGPGQQTSGRGGLGG 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QGAGAAAA------AAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAA- 234
GQGAGAAAAAAAAAAG-GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGA--- 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     44 GGYG----GLGSQGTSGRG----- 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 ASMTGGQQMGRGSMASGRGGLGGQGAGAAA-------AAAAAAGGAGQ 43
                                                                                               GQGGYG---GLGSQG-----TSGPGGYGPGQQTSG-RGGLGGQGAGAAAAAAA
                                                                                                                                                                                                                                     581 GOGGYGNORGVGSYGOGAGAGAATSAAGGAGSGRGGYGEOGGLGGYGOGAGAGAASTAA
                                                                                                                                                                                                                                                                                                       587 RGCLGGQGAGAAAAAAAAG-GAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAG
                                                                                                                                                                                                     531 GGA---GQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGG-AGQGGYGGLGSQGTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GAGQGGYG----GLGSQGTSGRG------GLGGQGAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Araneae;
Nephila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21; Mismatches 184; Indels 473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 5; Length 1953;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE-21179804; PubMed=11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Extreme Diversity, Conservation, and Convergence of Spider Fibroin Sequences.";
Science 291:2603-2605(2001).
EMBL; AF350276; AAK30605.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 1953 1953
1953 AA: 159383 MW; C82B4DACD043C3BD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Major ampullate spidroin 2-like protein (Fragment).
Nephila madagascariensis.
Eukaryota, Metazoa, Arthropoda, Chelicerata, Arachnida,
Araneomorphae, Entelegynae, Araneoidea, Tetragnathidae,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41.0%; Score 1452.5; DE 39.6%; Pred. No. 3.2e-71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 1953 AA
                                                                                                                                                                                                                                                                                                                                                                                                             646 GAGQGGYGGLGSQGTSGPGQQTSGIRI 677
                                                                                                                                                                                                                                                                                                                                                                                                                                                  747 GAG-GGGGGGGGGGGGGGGGGSSAGLSV 777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 3.00 Matches 445; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=115969;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9BIT7
                                                                                                  484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  116
                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SETT REPAIRS SO SET THE SET TH
                                                                                                  à
                                                                                                                                               g
                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
```

MEDLINE-98200471; PubMed-9541398;
Colgin M.A., Lewis R.V.;
"Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like 'spacer regions'.";
Protein Sci. 7:667-672(1998). 01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-JUN-1998 (TrEMBLrel. 17, Last annotation update)
Minor ampullate silk protein MiSpl (Fragment).
Nephlla clavipes (Orb spider).
Bukaryota; Metazoa; Arthropoda; Chelleerata; Arachnida;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae;

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                 339
                                                                                                                                                                                                                                                                                                                                                                                                384
                                                                                                                                                                                                                                                                                                                                                                                                                                                46
                                                                                                                                       53
                                                                                                                                                                                                                                    167
                                                                                                                                                                                                                                                                    213
                                                                                                                                                                                                                                                                                                                                                                                627
                                                                                                                                                                                                                                                                                                                                                                                                                989
                                                                                                                                                                                                                                                                                                                                                                                                                                421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              925
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              594
    g
                                                                                                                                       ò
                                                                                                                                                      g
                                                                                                                                                                                     8
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                         ò
                                                                                                                                                                      ò
                                                                                                                                                                                                     ò
                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                 25;
                                                                                         AAGAGAGGYGGQGGYGAGAGAGAAAAAGAGAGGAGGYGRGAGAGAAA--GAGAGGYGG 600
                                                                                                              GYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYG----GLGSQGTSGRGGLG 137
                                                                                                                         GQGAGAAAAAAAAAAGGAGGGYGGLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAA 197
                                                                                                                                                       257
                                                                                                                                                                                                            GRGCLGGQGAGAAAAA-AAAAAGGAGQGGYGGLG-----SQGTSGRGCLGGQGAGAA 308
                                                                                                                                                                                                                         AGGAGQGGYGGLGSQGT-----SGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG 414
                                                                                                                                                                                                                                                                                                         LGSQG------TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGG 463
                                                                                                                                                                                                                                                                                                                          651
                                                                                                                                                                                                                                                                                                                                         LGGQGAGAAAAAAAAAAGGAGQGGYG---GLGSQGTSGPGGYGPGQQTSGRGGLGGQGA- 519
                                                                                                                                                                                                                                                                                                                                                  -GLGS-----QGTSGRGGLG-GQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 630
                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Antheraea pernyi (Chinese oak silk moth).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Saturniidae; Saturniinae; Saturniini; Antheraea.
[1]
                                                                               Indels 117;
                                                Length 988;
                                988 AA; 79082 MW; 461E03DF53F7085D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Pred. No. 2.6e-71; 17; Mismatches 212;
                                                DB 5;
                                                40.9%; Score 1450; DB 5
50.6%; Pred. No. 2.6e-71
                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 2639 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last seq
01-DEC-2001 (TrEMBLrel. 19, Last anno
EMBL; AF027735; AAC14589.1; -. InterPro; IPR000817; Prion. PRINTS; PR00341; PRION. NON_TER 1 1 SEQUENCE 988 AA; 79082 MW;
                                                        Best Local Similarity 50.6%
Watches 354; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Fibroin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     076786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            076786
                                                                               23
                                                                                                                                                            325
                                                                                                                                                                                                                                                                         366
                                                                                                                                                                                                                                                                                                                                                                                                                      694
                                                                                                                                                                                                                                                                                                                                                                                                                                      631
                                                                                              20
                                                                                                              83
                                                                                                                              271
                                                                                                                                              138
                                                                                                                                                                            198
                                                                                                                                                                                            383
                                                                                                                                                                                                                           423
                                                                                                                                                                                                                                           309
                                                                                                                                                                                                                                                           483
                                                                                                                                                                                                                                                                                         543
                                                                                                                                                                                                                                                                                                         415
                                                                                                                                                                                                                                                                                                                        601
                                                                                                                                                                                                                                                                                                                                        464
                                                                                                                                                                                                                                                                                                                                                        652
                                                                                                                                                                                                                                                                                                                                                                                      712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0767
 ST EN
                                                                                             g
                                                                                                                            a
                                                                                                                                                                                           g
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                              ò
                                                                                                                                             à
                                                                                                                                                          셤
                                                                                                                                                                           ò
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            924
                                                                                                                          gene
                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      254 QGTSGRGGLGGQGAGAAAAAAA---AGGAQQGGYGGLGSQGT-----SGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          508 YGW-GDGGYGSDSAAAAAAAAAAAAAAGSGAGGRGDGGYGSGSSAAAAAAAAAAAAAARRAGH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------SGRGGLGGQCAGAAAAAAAAGGAG--QGGYGGLGSQGTSGRGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            105 AAAAAAAAAAAGGAGGGYGGLGSQGTS-----GRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       138 GQGAGAAAAAAAA----AGGAGQGYGGLGSQ---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             || || || || || 391 AAAAAAAAAAAAAAAAAAGSGAGGAGGGYGWGDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------LGSQGTSGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGLGGQGAGAAAAAAAAAAGGAG--QGGY----GGLGSQGT-----SGPGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GRGGLGG-------QGAGAAAAAAA---AGGAGQGGYGGLGSQGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GAGGVGGGYGWGDGGYGSDSAAAAAAAAAAAAGSGAGGRGDGGYGSGSSAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----GISGPGGYGPGQQISGRGGLGGGGAGAAAAAAAAAAGGAGGGGYGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GPGQQTSGRGGLGGQGAGAAAAAAAA--GGAGQGGYGGLGSQGT-----S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----SGRGGLGGQGAGAAAAAAAAAGGAG--QGGY----GGLGSQGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         459 -SGRGGLGG------QGAGAAAAAAAAAAGGAGQGGYGGLGSQGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----SGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAA----AGGAGQGGYGGLGSQGT-
                                                                                                                          flbroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 2639;
SEQUENCE FROM N.A.

MEDILINE-20496933; Pubmed-11040284;
Sezutsu H., Yukuhico K.; Elizabetsu H., Four types of repetitive units.";
J. Mol. Evol. 51:329-338 (2000).

EMBI, AF083334; AAC12606.1; Elizabetsu Eliz
                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                               39.1%; Score 1386; DB 5; I
40.6%; Pred. No. 1.6e-67;
tive 23; Mismatches 248;
                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
Matches 400; Conserv
```

^

ò	628 1.GGOCAGAAAAAAAAAGGAGGGGYGG1.GSO	è	327
7 A		3 2	1809 WGDGGYGSDSAAA
ò	PGGYGPGQQTSGIRIRAPSTSFEHHH 688	ò	
og o	: :	7 d	
1000	Ç.	δλ	414
0964 1D	RESULT IO Q964F4 PRELIMINARY; PRT; 2655 AA.	QQ	1920 AAAARRAGHDRAAG
AC T	; -2001 (TrEMBLre) 19. Created)	οy	442 AGGAGOGGYGGLGS
占占	(TrEMBLrel. 19, (TrEMBLrel. 19,	qq	1980 AAGSGAGGVGG
OS	Fibroin. Antheraea yamamai.	oy o	493 SQ
88	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;	Q	2036 SDSAAAAAAAAA
0 8 8 8 8 8	Bombycoidea; Saturnildae; Saturniinae; Saturniini; Antheraea. NCBL_TaxID=7121; [11]	Š ć	523 AAAAAAAAGGGG-
R P	TRANSPOSON-MARINER-LIKE ELEMENT;	a vo	553 GGQGAGAAAAAA
8 8 E	Lee JS., Hwang JS., Kang SW., Suh DS., Jin BR., Kim YS., Lee KS., Goo TW., Yun EY., Kim K.Y.;	g	
R F	"Complete nucleotide sequence of ilbroin gene cloned from Antheraea yamamai.";	٥	593QGAGAAAAA
2	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases. EMBL; AF125500; AR813145.1; SECHIENCE: 2655 as 216573 ww. RORRELDEFOCACA6 CPC64.	q ₀	 2216 GYGSDSAAAAAA
3	ACLUMN, GIOU'S MN, CONDUCTOR OCCUPACION OF THE CO.	٥٧	625 RGGLGGQGA-GAA
E B C	n 2655; 415. Cand	q	2276 DGGYGGYGSDSAAA
Ę ;	*COMPACTOR TANDAM TANDAM TO COMPACT THE TO COMPACT TO C	Qy	684 FEHH 687
<u>3</u> 8	2 ASMILOSQUEROSSMASOKO-GLOSQUARAKAKAKAKAKAKOKOGOKOKOKOKOKOKOKOKOLOSQUEROKKOKOLOSU 1	qa	2333 RAGH 2336
οy	61 GGQGAGAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGG 100	RESULT	RESULT 11
οg	ASGAGGA	900	Q8VIY9 PRELIM
οy	101QGAGAAAAAAAAAGGGGGGGGGGS	A E	01-MAR-2002 (TrEMBI
Op	1390 DGGYGSDSAAAAAAAAAAAASGAGGRGDGGYGSGSSAAAAAAAAAAAAAAAAAAAAARRAGYDRAAGS 1449	ដែដ	01-MAR-2002 (TrEMBI 01-JUN-2002 (TrEMBI
οy	GAGO	GE	MT3615.3
qq	1450 AAAAAAAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGG	8 O S	Mycobacterium tuber Bacteria; Firmicute
ΟY	165 SQCTSGPGGYGPGQCTSGRGGLGGQGAGAAAAAAAAAGGAGQ 207	2 8 2	Actinomycetales; CC NCBL_TaxID=1773;
qq	1510 SDSAAAAAAAAAAAAAGGGGGGGGGGGGGGGGGGGGGGG	RP C	SEQUENCE FROM N.A.
οy	208 GGYG 226	RA SA	Fleischmann R.D., Poton
qq	1569 GGYGSGSSAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	R R	Kolonay J.F., Nelsc
οy	227	R R F	Describer A., Occessor Bishai W.; "Whole denome compa
qq	BAGGA	RI	laboratory strains. Submitted (APR-2001
ò i		D R	EMBL; AE007164; AAR TIGR; MT3615; -
QQ	gdcgygsdsaaaaaaaaaagsgaggrgdggygsssaaaaaaaaaaarraghdraa	g g g	InterPro; IPR000084 InterPro; IPR002173
9 6	288326 	8 8 8	Pfam; PF00934; PE; ProDom; PD001223; I PROSITE; PS00583; I
		S	SEQUENCE 1217 AA,

```
OSHKOSH;
Alland D., Eisen J.A., Carpenter L., White O.,
y R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
son W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
back T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GSAAAAAAAAAAAASGAGGSGGGYGWGDGGYGSNSAAAAAAAAA 1979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAAAASGAGGSGGYGGYGSDSAAAAAAAAAAAAAGSCAGGYGGGYGWG 2275
                                                                                                                                              -----GLGSQGTS---GRGGLGGQGAGAAAAAAA 441
                                                                                                                                                                                                                                                                                            -----GTSGPGGYGPGQ------QTSGRGGLGGQGAGAA 522
                                                                          SGRGGLGGQGAGAAAAAAAA----AGGAGQGGYG-------413
                                                                                                                                                                                                                      SSQGTS-GRGCLGGQGAGAAAAAAAAAGGAGQGGY-----GGLG 492
                                                                                                                                                                                                                                                                                                                                                                   2--GGYG-----GRGGL 552
                                                                                                                                                                                                                                                                                                                                                                                                                                             AAAAGGAGQGGYGGLGSQGT-----SGRGGLGG----- 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAAAAGGAGGGYGGLGSQGTS-----G 624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        arculosis.
tes; Actinobacteria; Actinobacteridae;
Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 parison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  s";
01) to the EMBL/GenBank/DDBJ databases.
AR47974.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; PE_region; 1.
; PFKB_KINASES_1; UNKNOWN_3.
AA; 94463 MW; 2FSBDCASA30A056E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 1217 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        34; PE_region.
73; PfkB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Birel. 20, C
Birel. 20, I
Birel. 21, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        otein.
```

ω

```
36
                                                                                                                                                                                                                                                                                                                                                                                 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                               188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     762
                                                                                                                                                                                                                                                                                                                                                                                                                        137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         485
                                                                                                                                                                                                                                                                                                                                                                                                                                           407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 425
g
                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                 ģ
                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                            35;
                                                                214 GGGAGGVGGAGGGTGGAGGRAELLFGAGGAGGAGTDGGPGATGGTGGHGGVGGDGGWLAP 273
                                                                                                                                                                                       451
                                                                                                                                                                                                                                                            313
                                                                                                                                                                                                                                                                                                                               357
                                                                                                                                                                                                                                                                                                                                                                                                              442
                                                                                                                                                                                                                                                                                                                                                                                                                                 808
                                                                                                                                                                                                                                                                                                                                                                                                                                                     497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   609
                                                                                                                                                                   TSGRGGLGGQGA-------GQGGYG 161
                            Gaps
                                               ----- AA 35
                                                                                               ---GTSGRGGLGGQ----GAGAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                        GLGSQGTSGPGGYGPGQQTSGRGGLGGQGAGAAAAAAAAAAAAAAGGGGYGG-LGSQGTSG
                                                                                                                                                                                                                                                                                        GQGAGAAAAAA-----AAAGGAGQGGYGGLGSQGT---SGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                          572 GOGGGAGGSSGAGGINGSGCAGGIGGQGGAGGAGGAGADNPTGIGGTGGDGGTGGAAGAG
                                                                                                                                                                                                                                                                                                                                                                                                                       --GDGGAGGAGADADQPGATGGTGFAGGAGGAGGSSGAGGTNGSGGAGGTNGSGGAG
                                                                                                                                                                                                                                               RGGLGGQGAGAAAAAAAA------AAGGAGQGGYGGL----GSQGTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                 632 GAGGAAGTGGTGGMIGTTGNAGVGGAG-GQ--GGDGGAGGAGADADQPGATGGTGFAGGA
                                                                                                                                                                                                                                                                                                                                                                     358 AAAAAAAAAGG-----AGQGGYGGLGSQGT---SGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                         GGAGGAGGAGGNSGAGGTNGSGGAGGTGGQGGAGGAGGAGADNPTGIGGTGGDGGTGGAA
                                                                                                                                                                                                                                                                                                                                                                                                            399 AAAAAAGGAGQGGYGG-LGSQGTSGRGGLGGQ----GAGAAAAAA------AAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGA-GQGGYGGL----GSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLGSQGTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                               GPGG-YGPGQQTSGRGGLGGQGAGAAAAAAAAAAGGA----GQGGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           36 AAAGGA-GQGGYGGLGSQG--TSGRGGLGGQ-GAGAAAAAAAAAAGGAGQGGYGGLGSQ-
                                                                                                                                                                                                                            GTGGGGGAGGAGADNPTGIGGTGGDGGTGGAAGAGGAAGTGGTGGMIGTTGNAG
                                                                                                                                                                                                                                                                                                                              314 AAAGGAGGGYGG-LGSQGTSGPGGYGPGQQTSGRGGLGGQGA------GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GDGGNAGTGAGDPGKGGTGGTG-GTGGSGGAGGSGGANFNGCTGGTGGTG--GTGGKGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   924 NTDGLSSATSGTGGTGGTGGTGGAGDDSAGGTGGTGGAGGNAGAGGLANTGGTAGNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        610 ---GQGGYGGLGSQGTSG-----RGGLGGQG----AGAAAAAAAAAAAGGAGQGGY
                            210;
       Length 1217;
                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
        DB 16;
                           37; Mismatches 279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ą
       Score 1273;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 1715
                                             6 GGQQMGRGSMASGRGGLGGQ----GAGAAAA-
                  Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1044 GGAGISFSNGSNGGTGGTGGVG 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGLG-----SQGTSGPGGYG 667
       35.9%;
39.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (TrEMBLrel. 20, C
(TrEMBLrel. 20, I
(TrEMBLrel. 21, I
                           336; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAR-2002 (
01-MAR-2002 (
01-JUN-2002 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q8VI20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OSIV80
                  Best Loca
Matches
                                                                                                                                                                                                                             452
                                                                                                                           91
                                                                                                                                                                                    394
                                                                                                                                                                                                         162
                                                                                                                                                                                                                                                                                       264
                                                                                                                                                                                                                                               221
                                                                                                                                                                                                                                                                                                                                                                                                                                                   143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8VIZ0
                                                                                                      셤
                                                                                                                                             셤
                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                 g
                                              à
                                                                셤
                                                                                     ô
                                                                                                                           õ
                                                                                                                                                                   ò
                                                                                                                                                                                                        å
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
```

```
36;
                                                                                                                                                                     SEQUENCE FROM N.A.
STRAIN-CDC 1551 / OSHKOSH;
STRAIN-CDC 1551 / OSHKOSH;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khourl H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                        Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBL_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                  "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      227 GGGAGGVGGAGGCTCGAGGRAELLFCAGGAGGAGTDGGPGATGGTGGHGGVGGDGGWLAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAAGGA-GQGGYGGLGSQG--TSGRGGLGGQ-GAGAAAAAAAAAAAGGAGQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TSGR------GGLGGQGAGAAAAAAAAAGGAGQ-----GGYGGLGSQGTSGRGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGQG-AGAAAAAAAAAGGAG-QGGYGGLGSQ-----GTSGPGGYGPGQQTSGRGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGAGAAGADAPASTGLTGGTGFAGGAGGVGGQGGNAIAGGINGSGGAGGTGGQGGAGGMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GQCAGAAA-----GSQGTSGRGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GSGADNASGIGADGGAGGIGGNAGAGGAGGAAGTGGTGGVVGAAGKAGIGGTGGQGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QGAGAAAAAAAAAGGG----AGQGGYGGLGSQ----GTSGRGGLGGQGAGAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GTSGPGGYGPGQQTSGRGGLGGQGA-------GAAAAAAAAAAAAAGGA-GQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLGGQGAGAAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GTNGSG-----GAGGAGGKGGTGGAGGSGADNPTGAGFAGGAGGTGGAAGAGGAGGATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QGGYGGL----GSQGTSGPGGYG--PGQQTSG-----RGGLGGQGAGAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAGGAG-QGGYGGLGSQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GINGSGGAGGIGGGGGAGGAGAADNPIGIGGAGGIGGIGGAAGGAGGAIGIGGIGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 16; Length 1715;
-61;
282; Indels 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .00583; PFKB_KINASES_1; UNKNOWN_4.
1715 AA; 133832 MW; E857B2A774C771B0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                         laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: AE007163 AAK47971.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   35.8%; Score 1268.5;
40.5%; Pred. No. 2.5e-
ive 33; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGQQMGRGSMASGRGGLGGQ----GAGAAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00934; PE; I.
ProDom; PD001223; PE_region; I.
PROSITE; PS00583; PFKB_KINASES_SEQUENCE 1715 AA. 132027.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR000084; PE_region.
InterPro; IPR002173; PfkB.
                                              Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 35.8 Best Local Similarity 40.5 Matches 343; Conservative
family protein.
                                                                                                                                                                                                                                                                                                                          Bishai W.;
```

g

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 053557
                                                                                                                                                                                                                                                   256
                                                                                               g
                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HD DTT ACCOOK NAME OF ACCOOK NAME OF ACCOOK NAME ACCOO
                                                   õ
                                                                                                                                                 ò
                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAINR-98295987; PubMed-9634230;

MEDLINE-98295987; PubMed-9634230;

Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

Gordon S.V., Ediglameier K., Gas S., Barry C.E. III, Tekaia F.,

Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

Badcock R., Deviln K., Feltwall T., Gentles S., Hamlin N., Holroyd S.,

Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

A Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

A Hornsby T., Jagels K., Krogh A., Malandream M.A., Rogers J.,

Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

Neuter S., Seeger K., Whitehead S., Barrell B.G.;

Neuter S., Stalon S., Mycobacterium tuberculosis from the complete genome sequence.";

Nature 393:377-544(1998).

Ruberculist, Rv5514;

Ruberculist, Rv5514;

RitherPro; IPR000084; PE. region.
                                                                                                                                                                                                                            <u>،</u>
                                                                                               GGYGGLGSQC----TSGRGGLGGQCA----GAAAAAAAAAAAAGGA-----GQGGYGGLG 618
                                                                                                                         619 SQGTSGRGGLGGQGAGAAA-----AAAAAAAGGAGQGGYGGLGSQGTSGPG----GYGP 668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               s:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----GTSGRGGLGG----QGAGAAAAAAAAAGGAGQGGYGGLGSQGTS 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GTGGAGGDGVLGGVGGTGGKGGVGGVGGVGGVGGAGQLFSASGAAGNAGVGGAGGGG-- 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GRGGLGGQGA------GAAAAAAAAAAGGA------GQGGYGGLGSQGTSGPGGYG 175
  214 GGGAGGVGGAGGTGGAGGRAELLFGAGGAGGAGTDGGPGATGGTGGHGGVGGDGGWLAP 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----AA 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               36 AAAGGA-GQGGYGGLGSQG--TSGRGGLGGQ-GAGAAAAAAAAAAGGAGQGGYGGLGSQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1489;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           35.6%; Score 1260.5; DB 16; Length 1489
38.5%; Pred. No. 6e-61;
ive 35; Mismatches 274; Indels 241;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1489 AA; 115754 MW; 6855CBA1C3CBAF3A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6 GCQQMGRGSMASGRGGLGCQ----GAGAAAAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Probom; PD001223; PE_region; 1.
PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1489 AA
GGAGQGGYGGLGSQGTSG---RGGLGGQ-GAGAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            053559;
01-JUN-1998 (TrEMBLrel. 06, Cr
01-JUN-1998 (TrEMBLrel. 06, La
01-JUN-2002 (TrEMBLrel. 21, La
PGRS-family protein.
RV3514 OR MTV023.21.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 344; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00934; PE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Complete proteome. SEQUENCE 1489 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                1056 GQGGQG 1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                  GOOTSG 674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
053559
                                              876
                                                                                               574
                                                                                                                                              936
                                                                                                                                                                                                                                                                                               699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  91
                                                                                               ò
                                                                                                                                           원
                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          පු
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
STRAINHEATRY;

WEDLINE-98255987; PubMed-9634230;

Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekeis F.,

Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

Oliver S., Osborne J., Quall M.A., Rajandream M.A., Rogers J.,

Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   404 AGGAGQGGYGGLGSQGT---SGRGGLGGQGAGAAAAAAAAAAGGAGGGYGG-LGSQGTS 459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  227
                                                                                                                                                                                                                                                                                                                                                                                                           345 SGRGGLGGQGAGAAAAAAAAAGGAGQGGYGG-LGSQGTSGRGGLGGQGAGAAAAAAA 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     871 CGOVVAGGAGISFSNGSNGGTGGTGGVG--GTGGDGGNAGTGAGDPGKGGTGGTG-GTGG 927
392 GDGGAGGAGADADQPGATGGTGFAGGAGGAGGAGGSSGAGGTNGSGGAGGQGGAGGAGGA 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinomacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 176 PGQQTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGG-LGSQGTSGRGGLGGQ-----
                                                                                                                                                                                                                                                                                          512 GAGADADQPGATGGTGFAGGAGGAGGAGGSSGAGGTNGSGGAGGTGGQGGAGGAGAD
                                                                                                                                                                                                                                                                                                                                                                 -TSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGG-LGSQGTSGRGGLGGQ----GAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            310 AAA-----AAAAGGAGQ-----GGYGGLGSQ-GTSGPGGYGPGQQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                532 ADADQPGATGGTGFAGGAGGAGGAGGSSSAGGTNSSGSAGGTGRQSGTGGAGGAGADNPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 811 GVGGAGGQGGDGGAGGAGADADQPGATGGTGFAGGAGGAGGSGGSSCAGGTNGSGGAGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 484 --------GQGGYGGLGSQGTSGPGG-YGPGQQTSGRGCLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                524 AAAAAAGGGG-QGGYGGLGSQGTSGRGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       583 GTSGRGGLGGQGAGAAAAAAAAAGGA-----GQGGYGGLGSQGTSGRGGLGGQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1045 GAGGKGGAGGSSGAGGTNGSGGAGGAGGAGGAGISFSNGSNGGTGGTGGVG 1098
                                                                                                                                                                                                                      GAGAAA-----AAAAAAAGGA------GQGGYGGLGSQG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   632 -----GAGAAAAA----AAAAAGGAGQGGYGGLG-----SQGTSGPGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  460 GRGGLGGQ----GAGAAAAA------AAAAGGA-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
PGRS-family protein (Fragment).
RV3512 OR MTV023.19.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
```

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      147
        DDT HERE BY A PART HE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 470
                                                                                                                                                                                                                                                                                                                                            529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           618
                                                                                                                                                                                                                                                                                                                                                                                                    54 TSGRGGLGGQG-----AGAAAAAAAAAAAAGGA-GQGGYGGLGSQGTSGRGGLGGQ- 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----GAGAAAAAAAAGGAGGGYGG------LGSQGTSGRGGLGGQGAGAAA 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           176 GKGGDGGAGVTSSTAGNSGGAGGSGGKGGDAGAGGAGATPGANGIAGNGGDGGD--GAAG 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             234 AVGISGATGAGDGGHGGTGAAGGNGGTGGAGGSGIDGVGGGTGGTGGNGGNGAIGGAGGD 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QGGYGGLGSQGTSGRGGLGGQG------AGAAAAAAAAAAAAGGAG------QGGY 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGLGSQCTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQCTSGP---GGYGPGQQ 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    354 TGSAGGNG--GTGGRGGSGCAGGDGIGGVGGGKGCNGADGEVGGAGGAGGSGPNTSPGGN 411
                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                      ---LGSQG 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGGAG-QGGYGGLGSQ------GTSGRGGLGGQG--AGAAAAAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TGCDGGSGGAGGTGGAGGTGGAAGDGGQGGAGG-GAGGQGGAGGAGGTGGNGGNITG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGLGSQGTSGRGGLGGQGAGAAAAAAAAAGGAGQGGYGGLG----SQGTSGPGGYGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQTSGRGGLGGQGAGAA-AAAAAAGGAGQGGYG-GLGSQGTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GANGGAGGAGGSGGGTGGNGGAGGDAGDAGNGGNGTGNGGNGGNGGNGGNGGNGGAGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GSGNGGNGGNGGNAGNGGNSGTGSGDGGAGGNGGAACTGGTGGTGGTGGTGGTGGSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GAAAAAAAAAAAGG-AGQGGYGGLGSQ~-GTSGRGGLGGQGAGAAAAAAAAAAAGGAGQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGOG - - GNAGGGGAGGO - - GGADGGSGGDGGDAGTGGNGGNGGNRNSGNGTGGAGGNGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       564 AAAAAGGAGGGYGGL----GSQGT-SGRGGLGGQGAGAAAAAAAAAAGGAGGGYGGLG
"Deciphering the blology of Mycobacterium tuberculosis from the complete genome sequence.";
Nature 393:537-544(1998).
EMBL: AL022022; CAA1749.1; -.
Tuberculist; Rv3512; -.
InterPro; IPR002202; HMG-COA_red.
PROSITE; PS00318; HMG_COA_REDUCTASE_2; UNKNOWN_1.
                                                                                                                                                                                                                                                                             Indels 120;
                                                                                                                                                                                                                                    Length 1079;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --GTGGDGGNGGNGADNTANMTAQAGGDGGNGGDGGFGGGAGAGGGGGTTAGANGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SQGTSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGPGGYGPGQQTSG
                                                                                                                                                                                                                                                                                                                  TGGQQMGRGSMASGRGGLGGQGAGAAAAAAAAAAGGA-GQGGYGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TSGRGGLGGQGAGAAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQG--
                                                                                                                                                                          1
81163 MW; A79718CDCB74B97D CRC64;
                                                                                                                                                                                                                                35.1%; Score 1242; DB 16;
40.7%; Pred. No. 4.6e-60;
ive 34; Mismatches 306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
053775
ID 053775, PRELIMINARY; PRT;
AC 053775,
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
                                                                                                                                                                                                                                                   Best Local Similarity 40.74
Matches 316; Conservative
                                                                                                                                                                                         1079 AA;
                                                                                                                                              Complete proteome.
                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                      ហ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                994
    õ
                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ç° g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
```

```
<u>`</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67 AAAAAAAAAAGG-AGQGGYGG--LGSQGTSGRGGLGGQGA------GAAAAAAAAAAAAAAA 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
STRAIN-CDC 1551 / OSHKOSH;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Fleischmann T.D., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                         Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bishai W.; "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                      Harris
                                                                                                                                                                                                                                                                                                                                                    MEDLINE-9629587; PubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harriss Cole S.T., Brosch R., Parkhill J., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd Burnsby T., Jagels K., Kregh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandram M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; Complete genome sequence:

"Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11 GRGSMASGRGGLGGQGAGAAAAAAAAAAGG----AGQGGYGGLGSQGTSGRGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --AAAAAAGGAGGGYGGLGSQGTSG--PGGYG-PGQ--QTSGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAGG------AGQGGYGGLGSQ------GTSGRGGLGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGNGGNGGPGGAGGEGGAGGLALAASGONGSQGAGGDGGAGGNGGTPGNGGHGAAGALGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GG------AGQGGYGGL------GSQGTSGRGGLGGQGA------GAAAA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NGGVGGAGGHGGDPGVGGAGGQGGSGSTPGANGAPGNTPTSGGNGGNGGRGADATGFGQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          33.8%; Score 1198.5; DB 16; Length 1306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            363 363 G -> R (IN REF. 2).
1306 AA; 105964 MW; 843A30955FFA56B6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AL021942; CAAJ1449.1; -.
EMBL; AE006958; AAK44830.1; -.
01-JUN-1998 (TrEMBLrel. 06, Last sequence update) 01-JUN-2002 (TrEMBLrel. 21, Last annotation update) PGRS-family protein (PE_PGRS family protein). RV0578C OR MT0607 OR MTV039.16C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       37; Mismatches 293;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 1.2e-57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ProDom; P0001223; PE_region; 1.
PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_2.
PROSITE; PS00138; SUBTILASE_SER; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TubercuList; Rv0578c; ...
InterPro; IPR000209; Peptidase_S8.
InterPro; IPR000084; PE_region.
InterPro; IPR00173; PfkB.
Pfam; PF00934; PE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            38.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best_Local Similarity 38.4%
Matches 323; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Complete proteome.
                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TIGR; MT0607;
```

34;

```
619 SQCTS-GRGGLGGQGAGAAAAAAAAGGAGQGGYGGLG----SQCTSGPGGY-GPGQQT 672
           461 RGGLGGGGAGAAAAAAAAAAGGAGGGYGGLGSQGT-----SGPGGYGPGQQTS 509
                                                                                                                                                                                                                                        510 GRGGLGGQGAGAAAAAAAAAAAA-----GGAGQGGYGGLGSQGTSGRGGLGGQGA-G 558
                                                                                                                                                                                                                                                                                559 AAAAAAAAAAGGAGQGGYGGLGSQGTSGRGGLGGQGAGAAAAAAAAAAAAGGAGQGGYGGLG 618
235 AAAAAAGG-AGQGGYGGLGSQ-------GTSGRGGLGGQG--AGAAAAAAAA 277
                                                  GGYGP-----GQQTSGRGGLGGQG-----364
                                                                                                815 GGAGPAGTTGSQGAGGNGGSGGTGGDPGDGGNGANGSVFTNNGIGGNGGNGGNAGPSGAG 874
                                                                                                                  365 AAGGA------GQGGYGGLGSQGTSGRGCLGGQGAGAAAAAAAGGA 407
                                                                                                                                       875 GSGGAGSTFGATGSSSSIHVNGGNGGNGGNGDHALSGNGAAGGNGGNGGNGSLRGSGGAG 934
                                      AGGAGGGGGGGLGSQGTSGRGGLGGQGAGAAAAAAAAAG-GAGQGGYGGLGSQ-GTSGP
                                                                                                                                                                                                                                                                                                                                                           673 SG 674
:1
1228 AG 1229
                                      278
                                                                             336
                                  ŏ
                                                    ç
Q
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                            ò
```

Search completed: December 18, 2002, 16:20:53 Job time : 50.5691 secs This Page Blank (uspto)

Title: Perfect score:

Sequence:

•

OM protein

Run on:

Scoring table:

Searched:

DB seq DB seq

Minimum I Maximum I

Database

```
Dragline protein 1
N. clavipes spidro
N. clavipes spidro
N. clavipes spidro
Dragline protein 1
Dragline protein 1
Dragline protein 1
Spider dragline va
Polymer of an anal
Spider dragline va
Polymer of an anal
Spider dragline va
Polymer of an anal
Spider recombinant
Spider recombinant
N. clavipes spidro
N. clavipes spidro
Protein related to
Protein relates spidro
N. clavipes spider
Spider recombinant
Peptide dendrimer
Siprii protein com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sipili amino acid
Repetitive protein
SLPIII protein seq
Silk-like protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Silk like protein
Silk-like protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Spider; orb-weaver; sllk protein; pETNcDS; structural protein; purification; fibre; spinning.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Spider recombinant silk protein pETNcDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note- "encoded by TAT"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "encoded by CGA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note= "encoded by GAG
                                        AAMS004 2
AAU11794
AAU11794
AAU11794
AAR99055
AAY40101
AAR99057
AAY40101
AAR99057
AAW5004
AAM50045
AAM50045
AAM50045
AAM50045
AAM50045
AAM50045
AAM50045
AAM50045
AAM50031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB82611 standard; Protein; 528 AA
    01-NOV-2000; 2000WO-US30086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-JAN-2000; 2000US-0490291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          02-OCT-2001 (first entry)
  (MELL/) MELLO C M. (ARCI/) ARCIDIACONO S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nephila clavipes.
77722.0
77727.0
77707.1
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.0
77707.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         386.000003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO20015333-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   26-JUL-2001.
                                                                                            1914
1883.5
1883.5
1883.5
1883.5
1871.5
1858.5
1711.5
1711.5
1588.5
1588.5
                                                                                                                                                                                                                                                                                                                                                                   1234
1209.5
1144
                                                                                                                                                                                                                                                                                                                                                                                                                           1144
1144
1042.5
1008
973.5
972
                                                                                                                                                                                                                                                                                                                              1463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             972
972
972
972
972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB82611;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB82611
  Nephila clavipes s
Nephila clavipes s
N. Clavipes spider
N. Clavipes draglin
Spider silk protei
Spider natural sil
N. Clavipes spidro
N. Clavipes spidro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Spider recombinant
                                                                                                                                                    (without alignments)
1995.696 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
4. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
5. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
6. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
8. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
9. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
110. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
121. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
122. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
123. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
124. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
125. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
126. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
127. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
128. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
139. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
140. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
151. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
152. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
163. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
170. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
170. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
170. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
170. SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                                                                                                                                                                                                      US-09-490-291-8
2700
1 MASWIGGOQWGRIRIRGYGG......GLSGCDVLIQALLGHHHHH 528
                                                                                                                                December 18, 2002, 16:05:10 ; Search time 35.254 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /SIDS2/gogdata/geneseg/genesegp-embl/AA2000.DAT:*/SIDS2/gogdata/geneseg/genesegp-embl/AA2001.DAT:*/SIDS2/gogdata/geneseg/genesegp-embl/AA2002.DAT:*
                GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                           908470 seqs, 133250620 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                            protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW27178
AAW53346
AAX59070
AAX14308
AAX140097
AAM50037
AAM50039
AAM50038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB82611
                                                                                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A_Geneseq_101002:*
                                                                                                                                                                                                                                                                                                                                                                                                                       length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length DB
```

85.9 81.3 81.3 80.8 74.0 74.0 73.9 73.9

2319 2196 2196 2182 2182 1999 1999 1996 1996

12645078601

100.0

2700

Score

Result ş

421 GGQGVGAGGLGGQGAGAAAAVGAGQGGYGGVGSGASAASAAASRLSSPQASSRVSSAVSN 480

The present sequence is that of the orb-weaver spider (Nephila clavipes) recombinant silk protein pETNCDS. The invention provides for purifying and spinning spider silks and other structural proteins. Organic acids are used to lyse recombinant cells or other biological samples (such as non-recombinantly cells or other biological samples (such as non-recombinantly derived cells), and enrich the purity and yields of structural proteins intact. In the case of silk proteins, the structural proteins intact. In the case of silk proteins, the resulting lysate is further purified by ion-exchange or affinity of promappraphy and processed into an aqueous-based mixture for fibre spinning. In the present case, the pETNCDS gene was cloned into vector pET24 for recombinant expression in Escherichia coll, and recombinant silk protein was obtained using formic acid containing denaturant (3 M quanidine-HCI) and by affinity chromatography on nickel-NTA agarose. Products obtained using the method so of the invention can be used in the construction of many materials including films, fibres, woven articises, surfares, surfares, ballistic protection, parachtes and parachute cords. The new method has the following advantages over prior art: it involves center steps, requires less time and smaller volumes of reagents, results in better recovery of protein at higher purity (70-99%), is easy to scale up, and the fibres are spun in an environmentally benign solution reducing hazardous waste accumulation and cost. Recovering structural polypeptides in a biological sample, useful for purifying and spinning spider silks and other structural proteins, comprises treating the sample containing the polypeptides with an acid Butler MM; Claim 2; Page 41-42; 49pp; English. Arcidiacono S, (BUTL/) BUTLER M M. (USSA) US SEC OF ARMY. WPI; 2001-483136/52 N-PSDB; AAH26304 Mello CM,

528 AA; Sequence

ö 300 360 360 61 AGAAAAAAGGAGGGYGGLGSQGAGRGGLGGQGAGAAAAAGGVGQGGLGGQGAGQGAGAA 120 AAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLGG 180 Gaps 1 MASMTGGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQG 60 QGGLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAAGGAGQGGYGGQGAG QGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGGQGAGAAAAAAGGAGQGGLG GQGAGQGAGAAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAAAAAAGGAGQGGTGGL ö 100.0%; Score 2700; DB 22; Length 528; 100.0%; Pred. No. 1.6e-178; 1ve 0; Mismatches 0; Indels 0; Best Local Similarity 100. Matches 528; Conservative Query Match 61 121 121 301 181 241 301 361 ð g ò g ö 셤 ð g ò 셤 ò 셤 ò

10;

Gaps

Indels

Length 646;

85.9%; Score 2319; DB 18; 85.7%; Pred. No. 3.4e-152; Pred. No. 3.4e-152; 1; Mismatches 14;

Query Match 85.9 Best Local Similarity 85.7 Matches 484; Conservative

ò

361

g

6 GGQQMGRIRI-----RGYGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQG 60

encoding silk protein. The process involves: (a) selecting target DNA, crown a silk protein. The process involves: (a) selecting target DNA, from a silk-producing spider, that contains many repetitive and non-repetitive regions: (b) selecting a single-stranded DNA primer of at least 10 nucleotides with a sequence that is complementary to a region of the target; (c) repetitively combining the primer with melted target DNA, incubating the mixture with nucleotides and a DNA polymerase with prooffereding activity to produce a DNA fragment which is complementary to the target and is at least 2 kb long. The present sequence encodes the spider silk protein from Nephila clavipes. The DNA fragment can be used to make fibres, films, woven articles, e.g. for use in parachutes, sails, body armour, and absorbers (e.g. of heavy metals, biological weapons, DNA, chemicals, flavours and fragrances). The high molecular weight (90-250 kD) of spider silk proteins can be produced on a strength and elasticity than silkworm silk. Inclusion of both repetitive and non-repetitive regions ensures isolation of stable clones. New opt. multimerised DNA sequences encoding spider silk protein contg. both repetitive and non-repetitive sequences, useful for making high strength films, fibres, woven articles etc. High strength film; fibre; woven article; parachutes; sails; absorber; body armour; heavy metal; biological weapon; chemical; flavour; fragrance; Nephila clavipes. LVASGPINSAALSSTISNVVSQIGASNPGLSGCDVLIQALLGHHHHHH 528 Nephila clavipes spider silk protein. Ą Example 2; Fig 1; 57pp; English. AAW27178 standard; Protein; 646 96WO-US13767. 95US-0517694. (first entry) Basel RM, Elion GR; WPI; 1997-179272/16. (BASE/) BASEL R M. (ELIO/) ELION G R. 646 AA; Nephila clavipes. N-PSDB; AAT85356 WO9708315-A1. 22-AUG-1996; 22-AUG-1995; 09-DEC-1997 06-MAR-1997 AAW27178; 481 RESULT 2 AAW27178 셤 ò

m

```
The present sequence represents a spider silk protein from the present invention. Spider silk proteins, and peptide fragments within the proteins, can be produced and purified independently and can then be mixed and made into fibres that have higher tensile strengths and elasticity than naturally occurring fibres. The fibres can be used in mixed composites. The invention allows the two naturally occurring Nephila clavipes silk proteins to be produced independently so that they can later be combined to form silk fibres of high tensile strength and elasticity.
         Column 29-34; 68pp; English.
                                                                                                               718 AA;
                                                                                                               Sequence
                                                                                                                               Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY59070;
                                                                                                                                                                 ō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      486
                                                                                                                                                                                   22
                                                                                                                                                                                                    99
                                                                                                                                                                                                                   4
                                                                                                                                                                                                                                    120
                                                                                                                                                                                                                                                                                                                                          233
                                                                                                                                                                                                                                                                                                                                                                                            378
                                                                                                                                                                                                                                                                                                                                                                                                                                              337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      674
                                                                                                                                                                                                                                                                       167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY59070
ID AAY5
XX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
 ð
                                                                                                                                                                                 g
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                             Nephila clavipes; silk protein; tandem repeat; fibre; dragline; tensile strength; elasticity.
                                                   168
                                                                   238
                                                                                     228
                                                                                                     298
                                                                                                                      282
                                                                                                                               336
                                                                                                                                                                 376
                                                                                                                                                                                                   AAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAG 436
                                                                                                                                                                                                                                             522
                                                                                                                                                                                                                                                              496
                                                                                                                                                                                                                                                                               582
                                                                                             GAG--AAAAAGGAGGGLG----GQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAA
                                                                                                                                                       AAAAGGAGGGYGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAAAAGGAGQ-----
                                                                                                                                                                                        337 --------GGLGGQGAGAAAAAGGAGGGGLGGQGAGQGAGAAAAAA
                                                                                                                                                                                                                                   AAAAVGAGGGGGGGGGGGGASAASAASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTI
                                                                                    GSQGAGRGGLGGQGAGAAAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQ
                 AGAAAAAAG-------GAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQGGL
                         Recombinant spider silk proteins - useful for making fibres
                                                                                                                                                                                                                                                                                                                                                                                                            Nephila clavipes spider silk protein
                                                                                                                                                                                                                                                                                                                                                          Ź
                                                                                                                                                                                                                                                                                                               607
                                                                                                                                                                                                                                                                                                                                                          718
                                                                                                                                                                                                                                                                                              SNVVSQIGASNPGLSGCDVLIQALL
                                                                                                                                                                                                                                                                                                      SNVVSQIGASNPGLSGCDVLIQALL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91US-0684819.
90US-0511792.
94US-0317844.
95US-0425069.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ë
                                                                                                                                                                                                                                                                                                                                                          Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95US-0425069
                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       п×
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hinman MB, Lewis RV,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (UYWY-) UNIV WYOMING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1998-270437/24
                                                                                                                                                                                                                                                                                                                                                          standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAV23249
                                                                                                                                                                                                                                                                                                                                                                                           06-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-APR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           L5-APR-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-APR-1990;
04-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US5728810-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        17-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-APR-1995
                                                                                                                                                                                                                                                                                                                                                         AAW53346
                                                                                                                                                                                                                                                                                                                                                                          AAW53346;
                                                                                                                                                                                                                                                                                                                                                                                                                             Spider;
                                                                                                                                                                                                                                                                                                                                                                                                                                       coccon;
9
                                                                                                    239
                                                                                                                     229
                                                                                                                                                                                                                           377
                                                                                                                                                                                                                                            463
                61
                                                                                                                                                       283
                                                                                                                                                                                                                                                                                                               583
                                                                                                                                                                                                         401
                                                                                                                                                                                                                                                             437
                                                                                                                                                                                                                                                                                              497
                                                                                                                                                                                                                                                                                                                                       AAW53346
g
                 ð
                               g
                                                  ŏ
                                                                  g
                                                                                   ò
                                                                                                  g
                                                                                                                     à
                                                                                                                                    8
                                                                                                                                                    õ
                                                                                                                                                                     a
                                                                                                                                                                                        ò
                                                                                                                                                                                                        g
                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                              å
```

```
15;
                                                                                                                                                                                                                                                              336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             425
                                                                                                                     GLGSQGAGRGGLGGQGAGA----- 185
                                                                                                                                                                                                                                                                                                                              324
                                                                                                                                                                                                                                                                                                                                                                                  497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          198 AAAVGAGQEGIRGQGAGQGGXGGLGSQGSGRGGLGGQGAGAAAAAGGAGQGGLGGQGAG 557
                                                                                                       AAAGGAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQGGLGG---QGAGQGA-GA 119
                                                                                                                                                          AAAAAGGAGGGYGGLGSQGAGRGGSGGQGAGAAAAAAG-------GAGQGGYG 166
                            Gapa
                                                   GGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAA 65
                                                                 199 GLGSQGAGRGGLGGQGAGAAAAAAGGAGQGGLGGQGAGQGAGASAAAAGGAGQGGYGGL
                                                                                                                                                                                                                                                                                                                                                                                                                        366 QGAGAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAAAGGAGQGGTGGLGGQGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              285 AAGGAGGGYGG---QGAGQG------GYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAAAGGAGQGGLG----GQGAGAAAAAGGAGQGGYGGLGSQGAGR---GGQGAGA-AAA
                           Indels 188;
 Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           · :
                         10;
Score 2196; DB 19;
Pred. No. 1.1e-143;
                         4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ę
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY59070 standard; Protein; 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             694
81.3%;
70.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTNSAALSSTISNVVSQIGAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PINSAALSSTISNVVSQIGAS
                         Conservative
         ыевт Local Similarity
Matches 479; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                       325 AA-----
```

NAME OF COLOR OF STREET OF

```
425
The spider silk protein contains a basic 34 amino acid repeat. The repeat itself contains 3 regions. The first comprises 0-9 amino acids with a sequence AGR(GGX)2. This region is not highly conserved. The second region has a sequence GAG(A)x which is highly conserved and is 8-10 amino acids long. The third segment is (GGX)5 and is 15 amino acids long and 1s very highly conserved. In most cases X is A,O, Y or L. Removal of the poly-(Ala) segments results in a silk having lower elasticity.
                                                           AAAVGAGQEGIRGQGAGQGGYGGLGSQGSGRGGLGGQGAGAAAAAGGAGAGGGGGGGGGGGG
                                                                                                    QGAGAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAAAAAAAAAAAA
                                                                                                                     GAGGLGGQGAGAAAVGAGQGGYGGVGSGASAASAASRLSSPQASSRVSSAVSNLVASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6 GGQQMGRIRIRGYGGLGGQGAGQAAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA encoding spider silk protein-1 and 2 and variants - isolated from Nephila clavipes, for prodn. of spider silk protein and fibres having desired characteristics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 718;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein superfibre; major ampullate silk; orb web spider.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 2182; DB 12;
Pred. No. 1e-142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4; Mismatches
                                                                                                                                                                                                                                                                                                                                AAR14308 standard; Protein; 718 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                      N.clavipes dragline silk protein-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 15; Page 23; 48pp; English.
                                                                                                                                                                                                                          506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       90US-0511792
                                                                                                                                                                                                                                        PINSAALSSTISNVVSQIGAS
                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 477; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (UYWY-) UNIV OF WYOMING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1991-312199/43.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     718 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nephilia clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAQ14183.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Xu M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-APR-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-APR-1990;
                                                                                                                                                                                                                                                                                                                                                                                          15-JAN-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23-OCT-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EP452925-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lewis RV,
                                                                                                                                                                                                                                                                                                                                                            AAR14308;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
            438
                                                                                                                                558
                                                                                                                                                               426
                                                                                                                                                                                          614
                                                                                                    366
                                                                                                                                                                                                                         486
                                                                                                                                                                                                                                                     674
                                          337
                                                                                                                                                                                                                                                                                                                  AAR14308
                                        ò
                                                                                                  à
                                                                                                                                셤
                                                                                                                                                               ð
                                                                                                                                                                                          g
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15;
                                                                   Spider silk protein; dragline silk protein; major ampullate gland; fiber.
                                                                                                                                                                                                                                                                                                                                                                                                                                               The invention provides isolated cDNA molecules coding for spider silk proteins. The spider silk proteins are characterized by repeating alpha and beta regions and optional variable regions. The DNA sequences are useful in the production of spider silk protein by recombinant DNA techniques. The recombinant spider silk proteins may be used for the production of fibers. The present sequence represents the spider silk protein 1, derived from the major ampullate gland of Nephila clavipes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66 AAAGGAGGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQGGLGG---QGAGQGA-GA 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAG------GAGQGGYG 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22 GGQGAGQ---GGYGGLGGQGAGQGAGAAAAAAGGGGGGGGGGGGGGGGGAGAAA 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6 GGQQMGRIRIRGYGGLGGQGAGQAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Isolated DNA, vector and transformed cell encoding for and useful the production of spider silk protein \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAGGAGQGGYGG----QGAGQG-------GYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 718;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----AAAAGGAGQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 2196; DB 21;
Pred. No. 1.1e-143;
4; Mismatches 10;
                                       N. clavipes spider silk protein 1.
                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 1; Fig 6A-D; 65pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLGSQGAGRGGLGGQGAGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81.3%;
70.3%;
                                                                                                                                                                                                                                                                                                 Lewis RV;
                                                                                                                                                                                          94US-0317844
                                                                                                                                                                                                                    91US-0684819
90US-0511792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 479; Conservative
                                                                                                                                                                                                                                                                 (UYWY-) UNIV WYOMING.
                                                                                                                                                                                                                                                                                                                          WPI; 2000-061225/05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                               Xu M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       718 AA;
                                                                                                Nephila clavipes.
                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAZ38195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A. - - - - -
                                                                                                                                                                                        04-OCT-1994;
                                                                                                                                                                                                                      15-APR-1991;
                                                                                                                                                                                                                                    20-APR-1990;
                                                                                                                             US5989894-A.
                                                                                                                                                          23-NOV-1999
                                                                                                                                                                                                                                                                                               Hinman MB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      325
```

8 6 8 6 8 6

9 6 6 6 6 6

ò

쉽

δ

15;

containing spider silk protein,

dermatological composition containir skin care, in make-up or sunscreens

00.

```
WPI; 1999-510729/43
                 Cosmetic
for hair
                                                                                                                                                                                                                                                                                                                                                                                                                             AAU11781
ID AAU
g
                                                                                                                                                                         ò
                                                                                                                                                                                    셤
                                                                                                                                                                                                ò
                                                                                                                                                                                                           g
                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                           ò
                          185
                                                               258
                                                                                                  284
                                                                                                       -----AAAAAGGAGQ------ 336
                                                                                                                                                        365
                                                                                                                                                                                               QGAGAAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGV 425
                                                                                285 AAGGAGGGYGG---QGAGQG-------GYGGLGSQGAGRGGLGGQGAGA 324
                                                                                                                                557
                                                                                                                                                                                GLGSQGAGRGGLGGQGAGA-----
                                                        337 -----GGLGGQGAGAAAAAGGAGQGGLGGQGAG
                                                                                                                                                                                                      ------AAAAAGGAGQGGYGGLGGQGAGQGGYGCLGSQGAGGCGCGCGGGGAGA
                                                                                                  AAAAGGAGQGGLG----GQGAGAAAAAAGGAGQGGYGGLGSQGAGR---GGQGAGA-AAA
                                                                                                                                                                                                                                                                                                                                    Spider silk protein; spidroine major 1; cosmetic; make-up; dermatological compositions; hair care; skin care; sunscreen; hormone; moisturizer; skin disorder; skin disorder.
                                                                                                                                                                                                                                                                                                                         Spider silk protein spidroine major 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                    Arrandeau JP;
                                                                                                                                                                                                                                                                                       Ä.
                                                                                                                                                                                                                                                                                       651
                                                                                                                                                                                                                                              PINSAALSSTISNVVSQIGAS 506
                                                                                                                                                                                                                                                   PINSALSSTISNVSQIGAS 694
                                                                                                                                                                                                                                                                                       protein;
                                                                                                                                                                                                                                                                                                                                                                                                98FR-0001614
                                                                                                                                                                                                                                                                                                                                                                                                           98FR-0001614
                                                                                                                                                                                                                                                                                                                                                                                                                                   Philippe M, Garson JC,
                                                                                                                                                                                                                                                                                                                                                                                                                       (OREA ) L'OREAL SA.
                                                                                                                                                                                                                                                                                      AAY40097 standard;
                                                                                                                                                                                                                                                                                                                                                             Nephila clavipes
                                                                                                                                                325 AA----
                                                                                                                                                                                                                                                                                                                                                                                                11-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                            11-FEB-1998;
                                                                                                                                                                                                                                                                                                              19-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                    13-AUG-1999
                                                                                                                                                                                                                                                                                                   AAY40097;
                                                                                                                                                                                               366
                                                                          186
                                                                                                                                                                                                                                             486
                                                   167
                                                                                                  233
                                                                                                                                                                                                                                                                                AAY4009
                                                                                                                                                                                                                                                                           RESULT
              QQ
                            ò
                                      g
                                                 ö
                                                            g
                                                                         õ
                                                                                     g
                                                                                                ò
                                                                                                          셤
                                                                                                                        à
                                                                                                                                  g
                                                                                                                                               ò
                                                                                                                                                          g
                                                                                                                                                                        á
                                                                                                                                                                                  셤
                                                                                                                                                                                              ò
                                                                                                                                                                                                          g
                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                            ò
```

```
15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120 AAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAG------GAGQGGYG 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAAAAAGGAGQGGYGGLGNQGAGRGGQGAAAAAGGAGGGGYGGLGSQGAGRGGQGAGAA 497
                                                                                                                                                                                                                                                                                                                                                                                          The present sequence represents the natural spider silk protein spidroine major 1. The protein Improves the moisturizing/softening action of the compositions. The protein, and its fragments are used in cosmetic or dermatological compositions. These compositions have use as hair or skin care products; and make-up or sunscreens. As the protein is a good, persistent film-formers on the skin of low surface density, it can be used for delivery of active agents that are generally difficult to administer, e.g. vitamins, hormones, moisturizers or agents for treating disorders of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   498 AAAVGAGQEGIRGQGAGQGGYGGLGSQGSGRGGLGGGGAGAAAAAAGGAGGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6 GGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                     66 AAAGGAGGGGYGGLGSQGAGRGGLGGGGAG--AAAAAAGGVGQGGLGG---QGAGQGA-GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------AAAAAAGGAGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAGGAGGGYGG---QGAGQG-------GYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  366 QGAGAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             167 GLGSQGAGRGGLGGQGAGA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 188;
                                                                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                                                                          Score 1999; DB 20;
Pred. No. 3.8e-130;
5; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GAGGLGGQGAGAAAAVGAGQGGYGGVGSGASAASAAAS 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAU11781 standard; protein; 651
                                                                                                                                                                                                                                                                                          74.0%;
68.2%;
Claim 3; Fig 1; 32pp; French
                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 435; Conserv
                                                                                                                                                                                                                                                     651 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA-----
                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  614
```

```
AAM50037;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       wounds,
                                                                                                                                                                               325
                                                                                                                                                                                                                                                                                                                                                              366 (
                                                                                                                                                                                                                                                                       337
                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                          g
                                                                                        ò
                                                                                                                                  g
                                                                                                                                                                               à
                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, flexible tie downs for certain electrical components, or as a biomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is one-weaving spider spidroinl, a natural silk the method of the invention.
                                                                                                                                                   Orb-weaving spider; spidroin 1; silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAGGAGGGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQGGLGG---QGAGQGA-GA 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAAGGAGQGGYGGLGSQGAGRGGLGQQAGAAAAAAAGGAGQGGYGGLGNQGAGRGGQGA 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAAAAGGAGGGYGGLGSQGAGRGGSGGQGAGAAAAAG------GAGQGGYG 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGQQMGRIRIRGYGGLGGQGAGQGAGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------AAAAAAGGAGGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 651;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 1999; DB 23;
Pred. No. 3.8e-130;
5; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLGSQGAGRGGLGGQGAGA------
                                                                                                           Spider natural silk protein Spidroin 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 69-71; 93pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      expression cassette to a green plant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (DUPO ) DU PONT DE NEMOURS & CO E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      74.0%;
68.2%;
                                                                                                                                                                                                                                                                                                                                                                                24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAY-2000; 2000US-206968P
                                                            26-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            651 AA;
                                                                                                                                                                                                                                              Nephila clavipes
                                                                                                                                                                                                                                                                                        WO200190389-A2.
                                                                                                                                                                                                                                                                                                                                    29-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Sim
Matches 435;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Yang JG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             981
```

g

à

à g g

ð

셤

à

ç

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly; (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (1i) as adhession surfaces for culturing cells; and (1ii) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Also they retain water solubility after long-term boiling in aqueous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      closing
                                                  324
                                                                                    437
                                                                                                                        336
                                                                                                                                                          497
                                                                                                                                                                                              365
                                                                                                                                                                                                                                  557
                                                                                                                                                                                                                                                                                                        613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOISOI.
QGAGAAAAAAAAAAGGVRQGGXGGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGV
                                                                                                                                                                                                                                                                                                      558 QGAG----AAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGGQGV
                                                  --- GYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                          -----GGCGGGAGAAAAAAGGAGGAGGAG
                                                                                                                                                                                                                                498 AAAVGAGQEGIRGQGAGQGGYGGLGSQGSGRGGLGGQGAGAAAAAAGGAGQGGLGGQGAG
                                                                                                                                                          438 AAAAAAGGAGGGYGGLGNQGAGRGGQGAAAAAAGGAGGGGYGGLGSQGAGRGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New DNA encoding synthetic spider silk protein, useful e.g. for c.wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N. clavipes spidroin synthetic homologue SO1SO1 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                          426 GAGGLGGQGAGAAAVGAGQGGYGGVGSGASAASAAAS 463
                                                                                                                                                                                                                                                                                                                                                           Guehrs K;
                                                                                                                          ------AAAAAGGAGO-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 22; Page 29-34; 88pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Grosse F,
                                                 285 AAGGAGQGGYGG---QGAGQG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAM50037 standard; Protein; 1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-MAR-2001; 2001DE-1013781.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09-JUN-2000; 2000DE-1028212, 24-OCT-2000; 2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2002-123561/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; ABL61038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DE10113781-A1.
                                                                                                                          AA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
```

```
Query Match
                                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                        Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                            62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  337
                                                                                                                                                                                                                                                                                                                                                                                                                                         146
  ò
                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                             15;
     precipitated at high salt concentration, they are easily extracted and purified. The modular construction of the invention facilitates incorporation of additional peptide-encoding sequences, e.g. to simplify purification or modulate solubility. This sequence represents the synthetic N. clavipes spidroin-1 homologue SOISOI described in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Spidroin; spider; silk; fibre; film; membrane; wound; filter; SO1SO1SO1.
                                                                                                                                                                                                                                                                                                                                                                    145
                                                                                                                                                 279
                                                                                                                                                                                                                                                                                                336
                                                                                                                                                                                                                                                                                                        GGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAGAAAV--GAGQ 445
                                                                                                                   164
                                                                                                                                                                                                                     762
                                                                                                                                                                                                                                    224
                                                                                                                                                                                                                                                   821
                                                                                                                                                                                                                                                                                 881
                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                           61
                                                                                                                                                                                                                                                                         GGQQMGRIRI----RGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGA
                                                                                                                                                                                                       --------GGOGAGAAAAAGGAGOGG
                                                                                                                                                                                                               YGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGG
                                                                                                                                                                                                                                                                                               GAAAAAAGGAGQGGY - - - GGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAGGAGQ
                                                                                                                                                                                                                                           -GGLGGQGAGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGS---------
                                                                                                                                                                                                                                                                LGGQGAG-AAAAAGGAGQGGLG----GQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGA
solutions and since they are also soluble in organic solvents but
                                                                                            Indels 112;
                                                                                                                                         GAAAAAAGGAGGGGGGGGGGGGGGGGGGGGGGAGAAAAGGVGQ------
                                                                              Length 1255;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             clavipes spidroin synthetic homologue SOISOISO1 protein.
                                                                            Score 1996; DB 23;
Pred. No. 1.1e-129;
3; Mismatches 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                    GGYGGVGS------GASAASAAA 462
                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM50039 standard; Protein; 1880
                                                                           ttch 73.9%; sal Similarity 75.6%; 427; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                             1255 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DE10113781-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-DEC-2001
                                               invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                               Sequence
                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM50039
                                                                                     Local
                                                                                     Best Loca
Matches
                                                                                                                                        62
                                                                                                                                                                       106
                                                                                                                                                                                                                     703
                                                                                                                                                                                                                                    165
                                                                                                                                                                                                                                                                 225
                                                                                                                                                                                                                                                                                               280
                                                                                                                                                                                                                                                                                                               882
                                                                                                                                                                                                                                                                                                                             337
                                                                                                                                                                                                                                                                                                                                                                                                         1060
                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 9
AAM50039
8888888888
                                                                                                           ð
                                                                                                                         g
                                                                                                                                        ò
                                                                                                                                                       g
                                                                                                                                                                      à
                                                                                                                                                                                     g
                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
This invention describes a novel DNA sequence, encoding a synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at repeating unit of a spidroin protein. The synthetic protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (11) as adhesion surfaces for culturing cells; and (11) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Also they retain water solubility after long-term boiling in aqueous colutions and since they are also soluble in organic solvents but profited at high acit concentration, they are easily extracted and purified. The modular construction of the invention facilitates unfortable to purification or modulate solubility. This sequences, e.g. to simplify contained in the contained or solubility. This sequence represents the contained or solubility. This sequence represents the contained or solubility.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               used
                                                                                                                                                                                                                                                                   New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6 GGQQMGRIRI----RGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       703 LGSQGAGRGGLGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LGGQGAG-AAAAAGGAGQGGLG----GQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GAAAAAAGGGGGGY---GGQGAGQGGYGGLGSQGAGGGGGGGGGGAGAAAAAAAGGAGO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                   Guehrs K;
                                                                                                                                                                                                                                                                                                                                                                        Claim 22; Page 38-45; 88pp; German
                                                                                                                                   Grosse F,
2000DE-1028212
2000DE-1053478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                   Conrad
                                                                                                                                                                                     WPI; 2002-123561/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1880 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                 N-PSDB; ABL61040.
                         24-OCT-2000;
09-JUN-2000;
                                                                                                                                Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               427;
                                                                                                                                                                                                                                                                                                                           proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
```

```
26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAU11797;
                     ٥
                                                                               62
                                                                                                                                                                                                   144
                                                                                                                                                                                                                                                             186
                                                                                                                                                                                                                                                                                                                                                                                                                                           295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
                                                엄
                                                                                                      g
                                                                                                                                       á
                                                                                                                                                                    g
                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                               ò
                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This invention describes a novel DNA sequence, encoding a synthetic spider silk protein, comprising modules, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (11) as adhesion surfaces for culturing cells; and artificial organs; (11) as adhesion surfaces for culturing cells; and (11) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). CC fibres with excellent mechanical properties (strength and elasticity). CC Also they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high sait concentration, they are easily extracted and comportation of additional peptide-encoding sequences, e.g. to simplify purification or modulate solubility. This sequence represents the construction of the invention facilitates as synthetic construction of the invention facilitates.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                               Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOISM12.
445
                                              GGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAGAAAAV~-GAGQ
                                                                                                                                                                                                                                                                                                    N. clavipes spidroin synthetic homologue SOISM12 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Guehrs K;
                                                                                                                         1060 GGYGGLGSQGAGRGGQGAGAAAA 1084
                                                                                                                                                                                                           AAM50038 standard; Protein; 989 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 22; Page 34-38; 88pp; German.
                                                                                                         GGYGGVGS-----GASAASAAA 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Grosse F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000DE-1028212
2000DE-1053478
                                                                                                                                                                                                                                                                                                                                                                                                                                                      21-MAR-2001; 2001DE-1013781
                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2002-123561/17.
N-PSDB; ABL61039.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            989 AA;
                                                                                                                                                                                                                                                                                                                                                                                           DE10113781-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24-OCT-2000;
                                                                                                                                                                                                                                                                     18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                        13-DEC-2001
                                                                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                          AAM50038;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                         446
                                              388
                                                                                                                                                                                  RESULT 10
                                                                                                                                                                                                AAM5003B
                 g
                                                                         Q
                                                                                                                                                                                                                              à
                                                                                                         ö
                                                                                                                                    g
```

15;

Gaps

Length 989; Indels 170;

Score 1970; DB 23; Pred. No. 5.5e-128; 3; Mismatches 23;

73.0%; ilarity 68.6%; Conservative

Query Match Best Local Similarity Matches 428; Conserv

```
316
                                                                                                               143
                                                                                                                                           376
                                                                                                                                                                                  107
                                                                                                                                                                                                                                                    437 YGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGG 496
                                                                                                                                                                                                                                                                                   253
                                                                                                                                                                                                                                                                                                                                                                     616
                                                                                                                                                                                                                                                                                                                                        294
                                                                                                                                                                                                                                                                                                                                                                                                 337
                                                                                                                                                                                                                                                                                                                                                                                                                          674
                                                                                                                                                                                                                                                                                                                                                                                                                                                      389 GLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAGAAAV--GAGQG 446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                       Silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flaxible tie down; electrical component; implantation; Dragline protein 1; DP-1B 8mer; His tag.
              GAAAAAAG-------GAGGGGYGGLGSQGAGRGGLGGQGAG-AAAAAGGVGQGG
                                                                                                                                      317 LGGQGAGQGAGAAAAAGGAGQGGYGGLGSQCACRGGLGGQGACAAAAAAGGAGQGGYGG
                                                                                                                                                                                                                                                                                199 YGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAG-AAAAAGGAGQGGLG----GQGAGAAA
GGQQMGRIRI----RGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGA
                                                                                                                                                                                                                                                                                                                                                        -----GQQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                              -----GSGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGA-----
                                                                                                                                                                                                                                                                                                                                      254 AAAGGAGQGGYGGLGSQCAGRGGAGAAAAAGGAGQGGY-----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dragline protein 1 analogue DP-1B/his tag 8mer.
                                                                                                           108 LGGQGAGQGAGAAAAAGGAGQGGYGGLGSQGAGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ||||:||
| || ||:|||
| 795 GYGGLGSQGAGGAAAAA || 818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GYGGVGS-----GASAASAA 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAU11797 standard; Protein; 818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (DUPO ) DU PONT DE NEMOURS & CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-MAY-2001; 2001WO-US16937.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-MAY-2000; 2000US-206968P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nephila clavipes
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200190389-A2.
```

```
g
                                                                                       ö
                                                                                                                 g
                                                                                                                                                                                             ð
                                       ð
                                                                                                                                     The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, flexible tie downs for certain electrical components, or as a biomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dragline protein (DP) 1 synthetic variant DPLB/his tag 8mer (i.e. 8 copies of DPLB with a His tag at the C-terminus) which is used as an SLP construct in the method of the
                                                                                                                                                                                                                                                                                                                                                                                          20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     374 AAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAG------GAGQGGTGGL 420
                                                                                                                                                                                                                                                                                                                                                                                                                                 AGAAAAAAG-----GAGQGGYGGLGSQGAGRGGLGGQGAG-AAAAAGGVGQ 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      106 GGLGGOGAGOGAGAAAAAGGAGOGGYGGLGSOGAGGRGGSGGOGAGAAAAAAG----- 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              543 AGGAGQGGLGSQGAGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQ 602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------GGLGGQGAG------QGAGAAAA 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     363 AAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQ 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       423 GAGRGGLGGQGAGAAAAAAGGAGQGGLGSQGAGQAGAAAAAAGGAGQGGYGGLGSQGA 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAG 662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGQGVGAGGLGGQGAGAAAAV---GAGQGGYG--GVGSGASAAAAA-----SRLSS 467
                                                           Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein expression cassette to a green plant -
                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                  6 GGQQMGRIRI-----RGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQG 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----AAAAAGGGGYGGLGGQGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GAGRGGLGGGGAG--AAAAAGGGGGGG----GQGAGAAAAAAGGAGGGGYGGLGSQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        303 GGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GRGGQGAGAAAAAAGGAGQGGYGG---QGAGQGGYGGLGSQGAGRGGLGGQGAGAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----AAAAAAGGAGGGL------GGQGAGAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     159 -----GAGQGGYGGLGSQGAGRGGLGGQGAGA------GAGQGGYGGLGSQGAGRGGLGGQGAGA-----
                                                                                                                                                                                                                                                                                                                                                                                          Indels 228;
                                                                                                                                                                                                                                                                                                                                                               Score 1944; DB 23;
Pred. No. 2.9e-126;
3; Mismatches 48;
                                                                                                                 Example 1; Page 83-86; 93pp; English.
                                                                                                                                                                                                                                                                                                                                                             Query Match 72.0%;
Best Local Similarity 61.3%;
Matches 442; Conservative 3
                                    WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                      818 AA;
                                                                                                                                                                                                                                                                                                                                         Seguence
            Yang JG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     357
à
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant containing an SLP expression cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for diving the expression of an SLP transgene and T is a S' terminator. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, flexible tie downs for certain electrical components, or as a biomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dragline protein (DP) i synthetic variant DPIB/his tag 16mer (i.e. 16 copies of DPIB with a His tag at the C-terminus) which is used as an SLP construct in the method of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20;
                                                                527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Producing silk-like proteins in a green plant, useful in fabrics and material construction, comprises providing a silk-like protein expression cassette to a green plant -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGQQMGRIRI-----RGYGGLGGQGAGAAAAAAAGGAGGGGYGGLGSQCAGRGGG 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation; bragline protein 1; DP-IB 16mer; His tag.
                                                                                                    723 GSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGS
                                                             468 PQASSRVSSAVSNLVASGPINSAALSSTISNVVSQIGASNPGLSGCDVLIQALLGHHHHH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 228;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1626;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                72.0%; Score 1944; DB 23; 61.3%; Pred. No. 5.5e-126; iive 3; Mismatches 48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dragline protein 1 analogue DP-1B/his tag 16mer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 1; Page 87-92; 93pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ы
Н
                                                                                                                                                                                                                                                                                                                                                                                                         AAU11798 standard; Protein; 1626 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (DUPO ) DU PONT DE NEMOURS & CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-MAY-2001; 2001WO-US16937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-MAY-2000; 2000US-206968P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 72.03
Best Local Similarity 61.33
Matches 442; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1626 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200190389-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29-NOV-2001
                                                                                                                                                                                                                                              817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAU11798;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Yang JG;
                                                                                                                                                                                    I
                                                                                                                                                                                                                                              н
                                                                                                                                                                                    528
                                                                                                                                                                                                                                              817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ဖ
                                                                                                                                                                                                                                                                                                                                               RESULT 12
                                                                                                                                                                                                                                                                                                                                                                            AAU11798
```

(IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.

Grosse F,

Conrad U,

Scheller J,

```
ò
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                     1410
                      1591 QGAGRGGQGAGAAAAAAG-------GAGQGGYGG------LGSHHH 1624
                                                                                                                                                           325
                                                                                                                                                                                       356
                                                                                                                                                                                                                                               420
935 GGQGAGAAAAAAGGAGQGGLGSQGAGQGAG-AAAAAAGGAGQGGYGGLGSQGAGRGGQG 993
                                                                                                                                                                                                                                                                         GGQGVGAGGLGGQGAGAAAAV---GAGQGGYG--GVGSGASAASAAA----SRLSS 467
                                                  AGAAAAAAG------GAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQ
                                                                                                  ------AAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQ
                                                                                                                               GAGRGGLGGQGAG - - AAAAAGGAGGGCLG - - - GQGAGAAAAAAGGAGQGGYGGLGSQGA
                                                                                                                                                          GRGGQGAGAAAAAAGGAGQGGYGG---QGAGQGGYGGLGSQGAGRGGLGGQGAGAA----
                                                                                                                                                                                      -------GGQGAGAAAAAAGGAGQGGL------GGQGAGAAAAAAGGAGQ
                                                                                                                                                                                                   1351 AGGAGQGGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQ
                                                                                                                                                                                                                 ------QGAGAGA------QGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                    468 PQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCDVLIQALLGHHHHH
                                                                                                                                                                                                                                            374 AAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAG-------GAGQGGTGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                         Spidroin; spider; silk; fibre; film; membrane; wound; filter; SOI
                                         106 GGLGGQGAGQAQAAAAAGGAGQGGYGGLGSQGAGRGGSGGQCACAAAAAG---
                                                                      N. clavipes spidroin synthetic homologue SO1 protein
                                                                                                                                                                                                                                                                                                                                                                                   ¥.
                                                                                                                                                                                                                                                                                                                                                                                 AAM50042 standard; Protein; 630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-MAR-2001; 2001DE-1013781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000DE-1028212.
2000DE-1053478.
                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DE10113781-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09-JUN-2000;
24-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                             18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3-DEC-2001.
                                                                                                                                                                                                                                                                                                                                             1625 H 1625
                                                                                                                                                                                                                                                                                                                                Н 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                AAM50042;
              61
                                                                                                  186
                                                                                                                                                                                                                                                                         421
                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
AAM50042
                                                                                                                                                                                      326
                                                                                                                                                                                                                 357
                                                                                                             qq
셤
                           셤
                                         ð
                                                      셤
                                                                     ò
                                                                                 g
                                                                                                 à
                                                                                                                              à
                                                                                                                                        유
                                                                                                                                                         ò
                                                                                                                                                                      셤
                                                                                                                                                                                     ò
                                                                                                                                                                                                g
                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                 ò
```

```
squentially arranged oligonucleotides, each comprising a synthetic sequentially arranged oligonucleotides, each comprising a group of a sequentially arranged oligonucleotides, each oligonucleotide encoding a repeating unit of a spidroin protein. The synthetic protein has at least 84% homology with the Nephila clavipes spidroin protein and is used to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and clinis as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Calso they retain water solubility after long-term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and incorporation of additional peptide-encoding sequences, e.g. to simplify the protein or modulate solubility. This sequence represents the concentration.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14;
                                                                                                      New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66 AAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAGGVGQGGLGGQGAGQGAGAAAAAAG 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        102 GAGQGGYGGLGSQGAGR---GGQGAGAGAAAAAAGGAGQGGYGGLGSQGAGGGGGGGGAG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6 GGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65 AAAGGAGQGGYGGLGSQGAGRGGLG-------GQGAGAAAAAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 186 AAAAAAGGAGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAG-AAAAAGGAGQGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  298 GAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGGGGLG----GQGAGAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             354 AGQGGLGG---QGAGQG--AGAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AGGAGQGGTGGLGGQGVGAGGLGGQGAGAAAV--GAGQGGYGGVGS------GASAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                245 G----GQGAGAAAAAAGCAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGY---GGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 630;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 71.1%; Score 1919; DB 23; Best Local Similarity 83.5%; Pred. No: 1.2e-124; Matches 404; Conservative 4; Mismatches 16;
                                                                                                                                                                                                                                                                                               This invention describes a novel DNA sequence,
                                                                                                                                                                                                                                 Claim 22; Page 48-50; 88pp; German.
WPI; 2002-123561/17
N-PSDB; ABL61043.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      630 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 459 SAAA 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nvention.
                                                                                                                                                                       proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  409
```

Page 11

14;

60; Gaps

Indels

Pred. No. 1.3e-124; 4; Mismatches 16;

4

Conservative

83.58;

```
Best Local Similarity
                              404;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAU11793;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yang JG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            409
                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             298
                                                                                                                      셤
                                                                                                                                                                            ò
                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                          용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ints lines the describes a move that sequence, encouring a synthaetic sequentially arranged oligonucleotides, each comprising a group of sequentially arranged oligonucleotides, each oligonucleotide encoding a sequentially arranged oligonucleotides, each oligonucleotide encoding carepeating unit of a spidroin protein. The synthetic protein has at to produce synthetic fibres, films and/or membranes, particularly: (1) for medical use, especially to close wounds and/or to support or cover artificial organs; (ii) as adhesion surfaces for culturing cells; and (iii) as filters. The synthetic proteins are very similar to native spider silk proteins; can be prepared on a large scale and can be spun to fibres with excellent mechanical properties (strength and elasticity). Also they retain water solubility after long term boiling in aqueous solutions and since they are also soluble in organic solvents but precipitated at high salt concentration, they are easily extracted and purification of additional peptide-encoding sequences, e.g. to simplify unification or modulate solubility. This sequence represents a construct composed of the LeB4 signal peptide, N. clavipes spidroin-I synthetic homologue SOL, a c-Myctag and an endoplasmic reticulum (ER)-retention signal described in the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New DNA encoding synthetic spider silk protein, useful e.g. for closing wounds, comprises modules that encode repeating units of spidroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This invention describes a novel DNA sequence, encoding a synthetic
                                                                                                                                                                                                                                                                                                                                               Spidroin; spider; silk; fibre; film; membrane; wound; filter; SO1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "Synthetic spidroin homologue SOI"
660..672
/note= "c-myc-tag"

    N. clavipes spidroin synthetic homologue SO1 protein #2.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "ER retention signal"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Grosse F, Guehrs K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1..28
/label = LeB4_signal_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                 .ocation/Qualifiers
                                                                                                                                             AAM50047 standard; Protein; 676 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 1; Fig 10A; 88pp; German
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21-MAR-2001; 2001DE-1013781.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000DE-1028212.
2000DE-1053478.
                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note= "c
673..676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29..659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conrad U,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2002-123561/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     676 AA;
:111
456 AAAA 459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DE10113781-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24 -OCT-2000;
                                                                                                                                                                                                                                             18-SEP-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Scheller J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                              Synthetic
                                                                                                                                                                                            AAM50047;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Domain
                                                                                                                 AAM50047
                    g
```

23; Length 676;

8

Score 1919;

Query Match

```
Orb-weaving spider; silk-like protein; SLP; transgenic plant; promoter; 5' terminator; fabric production; material construction; rope; surgical suture; flexible tie down; electrical component; implantation; pragline protein 1; DP-18 Bmer.
                                                                        185
                                                                                                                                 245
                                                         66 AAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAGGVGQGGLGGQGAGQGAGAAAAAAG 125
                                                                                                                                                                           244
                                                                                                                                                                                                                                                               305
                                                                                                                                                                                                                                                                                           353
                                                                                                                                                                                                                                                                                                                      363
                                                                                                                                                                                                                                                                                                                                                   408
                                                                                                                                                                                                                                                                                                                                                                               423
                                                                                                                                                                                                                                                                                                                                                                                                                           245 G----GQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAAAAAAGGAGOGGY---GGO 297
   65
               Producing silk-like proteins in a green plant, useful in fabrics and
                                                                                                                                                                                          GGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA
                                                                                                                126 GAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                         186 AAAAAAGGAGGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAG-AAAAAGGAGQGGL
                                                                                                                                                                                                                                                 GAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAGGAGQGGLG----GQGAGAAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                   AGOGGLGG---QGAGQG--AGAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                          AGGACQCCTGCLGCQCVGAGGLGCQGAGAAAV--GAGQGGYGGVGS-----GASAA
                                                                                                                                                                                                                                                                                                                                                                            364 AGGGGYGGLGSQGAGRGGLGGQGAGAAAAAGGAGGGGYGGLGSQGAGRGGQGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dragline protein 1 analogue DP-1B 8mer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ε.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAU11793 standard; protein; 809 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (DUPO ) DU PONT DE NEMOURS & CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-MAY-2001; 2001WO-US16937.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-MAY-2000; 2000US-206968P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2002-106209/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200190389-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  459 SAAA 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             484 AAAA 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                29-NOV-2001
```

```
17;
                                                                          The invention relates to a method of producing silk-like proteins (SLP) in a green plant by providing a green plant contains the expression cassette. The green plant contains the expression cassette P-SLP-T where P is a promoter for driving the expression of an SLP transgene and T is a 5' terminator. The method is useful for producing silks and silk-like proteins in green plants. The silks and silk-like proteins may be used in fabrics or in material construction, such as rope, surgical sutures, flexible tie downs for certain electrical components, or as a biomaterial for implantation. The method allows for more cost effective production of silk not obtained from natural or microbial sources. The present sequence is the silk protein Dragline protein (DP) I synthetic variant DPIB Bmar (i.e. 8 copies of DPIB) which is used as an SLP construct in the method of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          184 AGAAAAAAGGAGGGYGGLGSQGAGQGGYGGLGSQGAGGGGGGGGGGAGAAAAAAGGAGQ 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        219 GAGRGGLGGQGAG--AAAAAGGAGQGGLG----GQGAGAAAAAAGGAGQGGYGGLGSQGA 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GRGGQGAGAAAAAGGACQCCYGC---QCAGQGGYGGLGSQGAGRGGLGGQGAGA---- 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       900
                                                                                                                                                                                                                                                                                                                                                                                                                             AGAAAAAAG------GAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQ 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---GGLGGQGAG------QGAGAAAA 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    601 GGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGGGGLGSQGAGQGAG 660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                           6 GCQQMGRIRI-----RGYGGLGGQGAGQAGAAAAAAAAGGAGGGGYGGLGSQGAGRGGQG 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----AAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -------GGQGAGAAAAAGGAGQGGL------GGQGAGAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     374 AAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAG-------GAGQGGTGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----GAGOGGYGGLGSQGAGRGGLGGQGAGA------
                                                                                                                                                                                                                                                                                                                                                                              Indels 194;
 construction, comprises providing a silk-like protein
                                                                                                                                                                                                                                                                                                                                          1 70.9%; Score 1914; DB 23; Length 809; Similarity 65.7%; Pred. No. 3.3e-124; 25; Conservative 1; Mismatches 27; Indels 194,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGQGVGAGGLGGQGAGAAAAV---GAGQGGYG--GVGSGASAASAAA 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 74-76; 93pp; English.
                expression cassette to a green plant
                                                                                                                                                                                                                                                                                                             809 AA;
                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                          Matches 425;
                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     421
à
                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
```

Search completed: December 18, 2002, 16:17:24 Job time: 40.254 secs

APP1.

_

Run

```
Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 747;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 2273; DB 3;
Pred. No. 2.7e-160;
                                        US-08-477-509B-83
US-08-482-085B-83
US-09-482-085B-83
US-08-482-085B-83
US-08-475-411A-41
US-08-475-411A-41
US-08-476-029A-36
US-08-477-509B-88
US-08-477-509B-88
US-08-444-791A-89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: DOS
SOFTWARE: FastSED for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/034,177
FILING DATE: HEREWITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN FIBROUS PROTEIN
NUMBER OF SEQUENCES: 3
NUMBER OF SECUENCES: 3
AUDRESSEE: Incyte Pharmaceuticals, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3: Incyte Pharmaceuticals, Inc
3174 Porter Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0486 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 3, Application US/09034177
Patent No. 6127146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84.2%;
71.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 747 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: 650-845-4166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
1059
1059
11101
11101
11101
11332
11332
11038
1038
1038
766
766
979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (: GenBank
GI 1174414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear IMMEDIATE SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET: 3174 POI
CITY: Palo Alto
STATE: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LIBRARY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; CLONE: (
US-09-034-177-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 1
US-09-034-177-3
962.5
9962.5
9962.5
9962.5
923.5
929.5
922.5
922.5
921.5
921.5
921.5
  Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (without alignments)
1277.936 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 2, A Sequence 19, A Sequence 11, A Sequence 13, Sequence 21, A Sequence 21, A Sequence 2, A Sequence 2, A Sequence 4, A Sequence 4, A Sequence 31, Sequence 31, Sequence 31, Sequence 31, Sequence 64, Seque
                                                                                                                                                           2002, 16:14:14 ; Search time 12.1566 Seconds
                                                                                                                                                                                                                                                                      2700
1 MASWIGGOQMGRIRIRGYGG.....GLSGCDVLIQALLGHHHHH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Issued_Patents_AA:*
. /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
. /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
. /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
. /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
. /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
. /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
. /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                                             Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-034-177-3

US-08-425-069-2

US-08-556-978B-19

US-09-247-806-1

US-09-247-806-6

US-09-247-806-6

US-09-247-806-6

US-09-247-806-6

US-08-556-978B-21

US-08-256-978B-21

US-08-256-978B-21

US-08-256-978B-21

US-08-256-978B-21

US-08-425-069-4

US-08-477-509-64

US-08-177-509-64

US-08-177-509-64

US-08-177-509-64

US-08-477-609-64

US-08-477-609-64

US-08-477-609-64

US-08-477-609-64

US-08-477-609-64

US-08-477-609-64

US-08-477-609-64

US-08-478-058-64

US-08-478-058-64

US-08-478-058-61
                       GenCore version 5.1.3 Copyright (c) 1993 - 2002 Compu
                                                                                                                                                                                                                                                                                                                                                                                                                                                       of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                            262574 segs, 29422922 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                using sw model
                                                                                                                                                                                                                                                                                                                                            BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                   US-09-490-291-8
                                                                                                                                                           December 18,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length

    protein search,

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2196
2196
21196
1999
1883.5
1883.5
1873.5
1343.5
1343.5
1144
1144
1144
                                                                                                                                                                                                                                                                              score:
                                                                                                                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                            rotal number
                                                                                                                OM protein
                                                                                                                                                                                                                                                                                                   Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database
                                                                                                                                                             on
O
                                                                                                                                                                                                                                                                           Perfect
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ş.
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                  å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
  15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Xu, Ming
APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
                                                                                                                                                                                                                                                                                                                                                                                                                                                   485
                                                                           365
                                                                                                                                                                                                                                                                                                                                                                                                           425
                                                               AAAGGAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAAGGVGQGGLGG---QGAGQGA-GA 119
                                                                                                         AAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAG------GAGQGGYG 166
                                                                                                                                                  GLGSQGAGRGGLGGQGAGA------ 185
                                                                                                                                                                                                                                      284
                                                                                                                                                                                                                                               324
                                                                                                                                                                                                                                                                                         557
                                                                                                                                                                                                                                                                                                                                           Gaps
                       65
                                 614 GRGGLGGGGAGAAAAGGGGGGGGGGGGGGASAASAAASRLSSPQASSRVSSAVSNLVASG
                     6 GGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA
                                                                                                                                                              ------AAAAAAGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                              AAGGAGGGYGG----QGAGQG-------GYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                               -------GGLGGQGAGAAAAAGGAGQGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                         QGAGAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAAAAAAGGAGGTGGLGGQGV
                                                                                                                                                                                                                                                                                                                                                                                                                   GAGGLGGQGAGAAAAVGAGQGGYGGVGSGASAASAAASRLSSPQASSRVSSAVSNLVASG
                                                                                                                                                                                                                                                                                                                        AA-----AAAAAGGAGO--------
 10; Indels 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5728810th Washington Street
CITY: Falls Church
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PINSAALSSTISNVVSQIGASNPGLSGCDVLIQALL 521
 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 2, Application US/08425069 Patent No. 5728810
 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: Floppy disk
IBM PC compatible
 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy of COMPUTER: IBM PC COMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 22046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Virginia
494;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-425-069-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STATE:
 Matches
                                                               99
                                                                                                        120
                                                                                                                                                  167
                                                                                                                                                                                            186
                                                                                                                                                                                                                                                                                                 378
                                                                                                                                                                                                                                   233
                                                                                                                                                                                                                                                                              285
                                                                                                                                                                                                                                                                                                                      325
                                                                                                                                                                                                                                                                                                                                                              337
                                                                                                                                                                                                                                                                                                                                                                                                        366
                                                                                                                                                                                                                                                                                                                                                                                                                                                  426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 2
                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                        셤
                                                               ö
                                                                                  ద
                                                                                                        ò
                                                                                                                          셤
                                                                                                                                                ð
                                                                                                                                                                                           à
                                                                                                                                                                                                            g
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                      õ
```

```
66 AAAGGACQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVQQGGLGG----QGACQCA-GA 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           425
                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                       ------AAAAAGGAGQGGYGCLGGQGAGQGGYGCLGSQGAGRGCLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    233 AAAAGGAGGGLG----GQGAGAAAAAAGGAGQGGYGGLGSQGAGR---GGQGAGA-AAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            285 AAGGAGGGYGG---QGAGQG------GYGGLGSQGAGRGGLGGOGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ||
|Aaaaaaggagyggyggignogagrggogaaaaaggagogyggigsogagggagaaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6 GGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120 AAAAAGGAGGGYGGLGSQGAGRGGSGGQGAGAAAAAAG------GAGQGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --------GGLGGQGAGAAAAGGAGGGGGGGGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GAGGLGGQGAGAAAVGAGQGGYGGVGSGASAASAASRLSSPQASSRVSSAVSNLVASG
                                                                                                                                                                                                                                                                                                                    Length 718;
          SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                 81.3%; Score 2196; DB 1;
70.3%; Pred. No. 1.2e-154;
ive 4; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLGSQGAGRGGLGGQGAGA-----
                                                               CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MUTPHY Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-106P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
                                         US/08/425,069
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PTNSAALSSTISNVVSQIGAS 506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO: 2:
                                                  19-APR-1995
                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: amino acid
                                                                                                                                                                      TELEFAX: (703) 205-8050
                                                                                                                                                                                                                                                                                                                             Best Local Similarity 70.3
Matches 479; Conservative
                                                                                                                                                                                                                                                         TOPOLOGY: linear MOLECULE TYPE: protein
                                        APPLICATION NUMBER:
FILING DATE: 19-APR
                                                                                                                                                                                                                                                                                  US-08-425-069-2
                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  337
```

15;

```
325 AA----
                                                                                                                                                                                                                                                                                                                                                                                                         US-08-556-978B-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGIH:
                                                                                                                                                                                                                                                                                                                                       $ 29
                                                                a
                                                                                                                                                                    ò
                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                       ద
                                 à
                                                                                                   à
                                                                                                                                   셤
                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                       ò
       Sequence 2, Application US/08317844B
Fatent No. 598984
GENERAL INFORMATION:
APPLICANT: Xu, Ming
APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5989894th Washington Street
CITY: Falls Church
STREET: 301 No. 5989894th Washington Street
COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               66 AAAGGAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQGGLGG---QGAGQGA-GA 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120 AAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAG-------GAGQGGYG 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     199 GLGSQCAGRGGLGGQGAGAAAAAAAGGAGQGGLGGQGAGQGAGASAAAAGGAGQGGYGGL 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------AAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGA 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6 GGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10; Indels 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 718;
                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,844B
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81.3%; Score 2196; DB 2; 70.3%; Pred. No. 1.2e-154;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLGSQGAGRGGLGGQGAGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 718 amino acids amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 70.3
Matches 479; Conservative
                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
US-08-317-844B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-317-844B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
15;
                                                                                       ---AAAAAGGAGQ----- 336
                                                                                                                                                                           498 AAAVGAGQEGIRGQGAGGGYGGLGSQGSGRGGLGGQGAGAAAAAAGGAGQGGLGGQGAG 557
                                                                                                                                                                                                                                                                   366 QGAGAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAAGGAGQGGTGGLGGQGV 425
                                                                                                                                                                                                                                                                                                                                                        GAGGLGGQGAGAAAAVGAGQGGYGGVGSGASAASAASRLSSPQASSRVSSAVSNLVASG 485
                                                                                                                                                                                                                                                                                                                                                                               285 AAGGAGQGGYGG---QGAGQG--------GYGGLGSQGAGRGGLGGQGAGA 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 74.0%; Score 1999; DB 4; Length 651; 68.2%; Pred. No. 3.7e-140; tive 5; Mismatches 10; Indels 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 19, Application US/08556978B
Patent No. 6268169
GENERAL INFORMATION:
APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED TITLE OF INVENTION: SPIDER SILK ANALOGS NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY STREET: 1007 MARKET STREET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: UNITED STATES OF AMERICA ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CR-9389-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08/077,600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION UNBER: 08/077,60(
FILING DATE: JUNE 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION UNBER: 33,692
REFERENCE/DOCKET NUBBER: CR-9
TELECOMMUNICATION:
TELECOMMUNICATION:
TELEPHONE: 302-892-8112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PINSAALSSTISNVVSQIGAS 694
                                                                                                                                                                                                                                                                                                                                                                                                                                              486 PTNSAALSSTISNVVSQIGAS 506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: 302-773-0164 INFORMATION FOR SEQ ID NO: 19:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             651 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 74.0
Best Local Similarity 68.2
Matches 435; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; MOLECULE TYPE: protein US-08-556-978B-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY: WILMINTON STATE: DELAWARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRANDEDNESS:
```

9 78 185

317 284 377 324 437 336 497

```
QGAGAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGV 425
                                                  22 GGGGAGO---GGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                      AAAGGAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAAGGVGQGGLGG---QGAGQGA-GA
                                                                                                                                                                                               120 AAAAAGGAGGGYGGLGSQGAGRGGSGGQGAGAAAAAAG-------GAGQGGYG
                                                                                                                                                                                                                                                                                                  GLGSQGAGRGGLGGQGAGA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAAAGGAGQGGLG----GQGAGAAAAAAGGAGQGGYGGLGSQGAGR----GGQGAGA-AAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAGGAGGGYGG----QGAGQG-------GYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----GELGGQGAGAAAAAGGAGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1111 | 1111 | 11111 | 11111 | 11111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 |
                                                                                                                                                                                                                                                                                                                                                                                                 -------AAAAAAGGAGGYGGLGGQGAGQGCYGGLGSQGAGRGCLGGQCAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: HIMINTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED TITLE OF INVENTION: SPIDER SILK ANALOGS NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GAGGLGGQGAGAAAVGAGQGGYGGVGSGASAASAAAS 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             614 GRGGLGGQGAGAAAGGAGQGYGGYGGYGGYGSASAASAAA 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: MICROSOFT WORD FOR WINDOWS CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: DISKFTTE, 3.50 INCH
COMPUTER: IEM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US/08/556,978B FILING DATE: CLASSFICATION: 435
PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/077,600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-556-978B-23
; Sequence 23, Application US/08556978B
; Patent No. 6268169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19898
                                                                                                                                                                                                                                               139
                                                                                                                                                                                                                                                                                                                                                                                               186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            285
                                                                                                  99
                                                                                                                                                                                                                                                                                               167
                                                                                                                                                                                                                                                                                                                                                                                                                                            259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                               g
                                                                                                                                          g
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                          염
                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1, Application US/09247806

Patent No. 6280747

GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: ARRAUDEAU, Jean-Claude
APPLICANT: ARRAUDEAU, Jean-Plerre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REPERBERCE: 6388-0365-0
CURRENT PILING DATE: 1999-02-11
SARLIER RAPLICATION NUMBER: FR 98/01614
EARLIER FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTIN VEY: 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15;
                                                                                        66 AAAGGACQGGYGGLGSQGACRGGLGGQGAG--AAAAAGGYGQGGLGG---QGAGQGA-GA 119
                                                                                                                         120 AAAAAGGAGQGGYGGLGSQGAGRGGSGGGAAAAAAAG------GAGQGGYG 166
                                                                                                                                                                                                                   - 185
                                                                                                                                                                                                                                                                                                                                                                                       ------AAAAAAGAGGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGA 232
                                                                                                                                                                                                                                                                                                                                                                                                                                          317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAGGAGQGGYGG---QGAGQG------GYGGLGSQGAGRGGLGGQGAGA 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAAAGGAGGGLG----GQGAGAAAAAGGAGQGGYGGLGSQGAGR---GGQGAGA-AAA 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QGAGAAAAAAAAAAAAGGVRQGCYGGLGSQGAGRGGQGAGAAAAAAAGGAGQGGTGGLGGQGV 425
                         Gaps
                                                                                                                                                                                                                                                                                                                 ---------GGLGGQGAGAAAAAGGAGGGGGGGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    498 AAAVGAGQEGIRGQGAGQGGYGGLGSQGSGRGGLGGQGAGAAAAAGGAGQGGLGGQGAG
GGQQMGRIRIRGYGGLGGQCAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 651;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 1999; DB 4;
Pred. No. 3.7e-140;
5; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            426 GAGGLGGQGAGAAAVGAGQGGYGGVGSGASAASAAS 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        614 GRGGLGGQGAGAAAGGAGQGGYGGVGSGASAASAAAA 651
                                                                                                                                                                                                                                                                                          GLGSQGAGRGGLGGQGAGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 74.0%;
Best Local Similarity 68.2%;
Matches 435; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Nephila clavipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-247-806-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-247-806-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  366
ø
                                                                                                                                                                                                                                                                                          167
                                                                                                                                                                                                                                                                                                                                                                                       186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  337
ç
                                          셤
                                                                                          å
                                                                                                                                     g
                                                                                                                                                                                        ò
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                             ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

557

16;

```
GRGGQGAGAAAAAAGGAGQGGYGG---QGAGQGGYGGLGSQGAGRGGLGCQGAGAAAAA 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      377 GRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAG-AAAAA 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GEMERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: ARRAUDEAU, Jean-Pierre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0355-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER APPLICATION NUMBER: FR 98/01614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    383 RQGGYGGLGSQGAGRG------GQG 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   495 GOGGYGGLGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGQGGLGSQGAGQG 554
                                                                                                                                                                                                                                                                                              61 AGAAAAAAG------GAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQ
                                                                                                                                                                                                                                                                                                                                                                                                                              106 GGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GAGRGGLGGQGAG--AAAAAGGAGGGLG----GQGAGAAAAAAGGAGQGGYGGLGSQGA
                                                                                                                                                                                                                                                                              6 GGQQMGRIRI----RGYGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             159 -----GAGQGGYGGLGSQGAGRGGLGGQGAGA------GAGQGGYGGLGSQGAGAGA-----
                                                                                                                                                                                                                                                 Indels 151;
                                                                                                                                                                                                                 Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    402 AGAAAAAGGAGGGGTGGLGGQGVGAGGLGGQGAGAAAAVGAGQGGYGGVGS 453
                                                                                                                                                                                                                Score 1883.5; DB 4;
Pred. No. 1.1e-131;
3; Mismatches 26;
                CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER APPLICATION NUMBER: FR 98/01614
EARLIER FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PALENTIN Ver. 2.1
SEQ ID NO 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Sequence 6, Application US/09247806
; Patent No. 6280747
                                                                                                                                                                                                                69.8%;
69.6%;
     FILE REFERENCE: 6388-0365-0
                                                                                                                                                               ; ORGANISM: Nephila clavipes US-09-247-806-8
                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                  Best Local Sim
Matches 412;
                                                                                                                                  909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
US-09-247-806-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       273
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                             엄
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                              16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 8, Application US/09247806
Patent No. 6280747
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL TOWN:
GENERAL TOWN:
GENERAL TOWN:
GENERAL TOWN:
GENERAL TOWN:
GENERAL OF THILIPPE, MICHEL
APPLICANT: GENERATION:
APPLICANT: GENERAL OR DEPLACE
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 AGAAAAAAG----AAAAAGCGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQ 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          106 GGLGGQGAGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAG----- 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            317 GAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAGAAAAAAGGAGGGGYGGLGSQGA 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    273 GRGGQGAGAAAAAGGAGQGGYGG---QGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAA 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAGGAGGGLG----GQGAGAAAAAGGAGQGGLGG---QGAGQGAGAAAAAAAAAGGV 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------GQG 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  554
                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                              GAGRGGLGGQGAG--AAAAAGGAGQGGLG----GQGAGAAAAAAGGAGQGGYGGLGSQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6 GGQQMGRIRI-----RGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GAGQGGYGGLGSQGAGRGGLGGQGAGA-------GAGQGGYGGLGSQGAGRGGLGG
                                                                                                                                                                                                                                                                                                              26; Indels 151;
                                                                                                                                                                                                                                                                              Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGAAAAAAGGAGGGGTGGLGGQGVGAGGLGGQGAGAAAAVGAGQGGYGGVGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 4;
                                                                                                                                                                                                                                                                                              Pred. No. 1.1e-131;
                                                                                                                                                                                                                                                                              69.8%; Score 1883.5; 69.6%; Pred. No. 1.1e
                                                                                                                                                                                                                                                                                                              3; Mismatches
FILING DATE: JUNE 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9389-A
TELECHONDICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFRAX: 302-773-0164
                                                                                                                               INFORMATION FOR SEQ ID NO: 23: SEQUENCE CHARACTERISTICS:
                                                                                                                                                               LENGTH: 606 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                            Matches 412; Conservative
                                                                                                                                                                                               unknown
                                                                                                                                                                                                                          MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                383 ROGGYGGLGSOGAGRG--
                                                                                                                                                                                             STRANDEDNESS: unkr
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                              US-08-556-978B-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-247-806-8
                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                              Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                402
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
```

à g ò g ò g g ò g à ద à g ò g

ö

g à ద

ò

```
142 GGLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR---GGQGAGAAAAAAGGAGQGGY 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 AGAAAAAAG----AAAAAGGGVGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQ 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          221 GRGGLGGQGAG--AAAAAGGAGQGGLG----GQGAGAAAAAAGGAGQGGYGGLGSQGAGR 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGYGGLGSQGAGRG------GQGAG 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6 GGQQMGRIRI-----RGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQG 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23 GGQGAGAAAAAAAGGAGGGLGSQGAGGAG-AAAAAAGGAGGGGYGGLGSQGAGRGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   106 GGLGGQGAQGAGAAAAAGGAGQGCYGGLGSQGAGRGGSGGQCAGAAAAAG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     259 AAAGGAGOGGYGGLGSQGAGRGGQGAGAAAAAAGGAGGGYGGLGSGGGGYGGLGSQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                275 GGQGAGAAAAAGGAGQGGYGG---QGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GGAGQGGLG----GQGAGAAAAAAGGAGGGLGG---QGAGQGAGAAAAAAAAAGGVRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        497 GEYGELGSQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGSQGAGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           159 -----GAGQGGYGGLGSQGAGRGGLGGQGAGA-------GAGQGGYGGLGSQGAGAGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAGGAGGGYGGLG-------GOGAGOGGYGGLGSOGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          69.3%; Score 1871.5; DB 4; Length 604; 69.8%; Pred. No. 8.5e-131; tve 3; Mismatches 26; Indels 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             404 AAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAGAAAAVGAGQGGYGGVG 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
                                                      APPLICATION NUMBER: US/08/556,978B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,600
FILING DATE: JUNE 15, 1993
ATTORNEY,AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/POCKET NUMBER: CR-9389-A
TELECOMMUNICATION INFORMATION:
TELECHONE: 302-692-8112
TELECHONE: 302-773-0164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Sequence 21, Application US/08556978B
                                                                                                                                                                                                                                                                                  TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 69.8
Matches 411; Conservative
                           SOFTWARE: MICROSOFT WC CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                             STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: unknown
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-556-978B-63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
US-08-556-978B-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           332 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                 16;
                                                                                                                                                                                                                                                                                                                                           259 AAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQ 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   330 AAGGACQGGLG----GQGAGAAAAAGGAGQGGLGG---QGAGQGAGAAAAAAAAAGGV 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       401
                                                                                                                                                                                                                                                                                                                       AGAAAAAAG-----AGAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQ 105
                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                     106 GGLGGQGAGAGAAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                  273 GRGGQGAGAAAAAAGGAGQGGYGG---QGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              497 GOGGYGGLGSOGAGOGGYGGLGSOGAGRGGLGGOGAGAAAAAAGGAGOGGLGSOGAGOG
                                                                                                                                                                                                                                                6 GGQQMGRIRI-----RGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGOG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       159 -----GAGQGGYGGLGSQGAGRGCLGGQGAGA------GAGQGGYGGLGSQGAGRGCLGGQGAGA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RQGGYGGLGSQGAGRG------GQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      219 GAGRGGLGGQG--AGAAAAGGAGQGGLG----GQGAGAAAAAAAGGAGQGGYGGLGSQGA
                                                                                                                                                                      Score 1873.5; DB 4; Length 606;
Pred. No. 6e-131;
3; Mismatches 27; Indels 151;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 63, Application US/08556978B
Patent No. 6268169
GENERAL INFORMATION:
APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
TITLE OF INVENTION: SPIDER SILK ANALOGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DELAWARE
UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 19898
COMPUTER READABLE FORM:
MEDIUW TREE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
EARLIER FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SCOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 606
                                                                                                                                                                      Query Match 69.4%;
Best Local Similarity 69.4%;
Matches 410; Conservative
                                                                                                              ; ORGANISM: Nephila clavipes US-09-247-806-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CITY: WILMINTON STATE: DELAWARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-556-978B-63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                              TYPE: PRT
                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                       19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           557
```

17;

188

220 318 378 331

```
107 GLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRG------GSGGQGAGAAAA 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Claude
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: LAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: EN 99/01614
EARLIER APPLICATION NUMBER: FN 99/01614
SEALLIER FILLING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTIN VOY: 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   158 GLGSQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQGAGRGGQGAGAAAA 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AG-AAAAAAGGGGGLGGQGAGQGAGAAAAAAAAA-----AGGVRQGGYGGLGSQ 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6 GCQQMGRIRI-----GAGGGGGGAGGGAGAAAAAAG------GAGGGGYGG 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       193 GAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAG------G
                                                                  394 GAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAGAAAAV---GAGQGGYG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  239 AGQCGLGGQGAG-AAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAAAGGAGQGGYGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --QGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAA-AAG------GAGQGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 1858.5; DB 4
Pred. No. 7.7e-130;
5; Mismatches 31;
                                                                                                                                                                                                                                                                   ; Sequence 4, Application US/09247806
; Patent No. 6280747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 68.8%;
Best Local Similarity 73.8%;
Matches 409; Conservative !
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT
ORGANISM: Nephila clavipes
                                                                                                                                                                            566 QGAGQGAGAAAAA 579
                                                                                                                                           450 -GVGSGASAASAAA 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ.
                                                                                                  g
                                 셤
                                                                    ò
                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       49 LGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAAGGVGQG 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               107 GLGGQGAGQGAGAGAAAAAGGAGQGGYGGLGSQGAGRG------GSGGQGAGAAAAA 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGGAGQGGYGGLGSQGAGRGGLGGQGAGA---------AAAAAG 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              193 GAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAG------G 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6 GGOOMGRIRI-----GAGOGAGOGAGAAAAAAAGG------GAGOGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGQGGLGGQGAG-AAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 1858.5; DB 4;
Pred. No. 7.7e-130;
5; Mismatches 31;
                                                                                                                    ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
COMPRATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B
                             APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY
TITLE OF INVENTION: SPIDER SILK ANALOGS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 08/077,600
FILING DATE: JUNE 15, 1993.
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9389-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEPHONE: 302-773-0164
                                                                                                                                                                                             UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO: 21:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     68.8%;
73.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best_Local Similarity 73.8
Matches 409; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: unknown
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             amino acid
                                                                                                                                                        CITY: WILMINTON STATE: DELAWARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                            19898
                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     239
```

8 % 8 % ò

g o

ò

18;

œ

```
210 YGAGAGAAAAAAAGAGSGGAGGYGRGAGAGAAAAAGAGSYGGQGYGAGAGAAA 269
                                                                                                                                                           330 AGAAAGAGGAGGYGGQSGYGAGAGAAAAAGGAGGAGGYGRGAGAGAGAAAAAAA 389
                                                                                                                                                                                                                         449 YGAGAGAAAAAATGAGGAGGYGRGAGAAAGAAAAGAGTGGAGYGGQGGYGAGAGA 508
                                                                                                                                                                                                                                                                                                                                                                                 GLGGQGAGQGAGAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGT 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      418 GGLGGQGVGAGGLGGQ-----GAGAAAAVGAGQGGYGGVGSGA----SAASAASRL 465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 683 ----GRGAGAGGYGGQGGYGAGAGAAAAGAGAGGYGDKEIACWSRCRYTVASTTSRL 738
                                              ---SQGAGRGGSG--GQGAGAAAAAAGGAGQGGYGGLG----
                                                                                248 -----GAGAAAAAAGGA-GQGGYG-----GLGSQGAGRGGQ-----GAGA
                                                                                                                                                                                                                                                                                                                                                        282 AAAAAGGAGQGGYG-----------GQGAGQGGYGGLGSQGAG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----SQGAGRGGLGGQ-----GAGAAAAAG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAG---GAGQGGLGGQ--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCDVLIQALL 521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
TITLE OF INVENTION: CDNAS Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US/08/458,298
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,747
FILING DATE: 14-MAR-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1447-104P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 2, Application US/08458298 Patent No. 5756677 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Murphy Jr., Gerald M. REGISTRATION NUMBER: 28 977 REEFENCE/DOCKET NUMBER: TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: USA
ZIP: 22040-3487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-458-298-2
                                                                                                                                                                                                 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     466
                                                                              g
                                                                                                                                                       g
                                            ò
                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25;
394 GAGRGGGGGAGAAAAAGGAGQGGTGGLGGGGGGGGGGGGGAAAV---GAGQGGYG- 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 49.8%; Score 1343.5; DB 1; Length 832; Best Local Similarity 47.9%; Pred. No. 9.4e-92; Matches 343; Conservative 19; Mismatches 143; Indels 211; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    49 LGSQCAGRGGQ----GAGAAAAAGGAGQGGYGGLG-----SQGAGRGGLGGQ--
                                      506 GAGRGGOGAGAAAAAGGAGOGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGQGGLGS
                                                                                                                                                                                                           Sequence 2, Application US/08209747
Patent No. 573371
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider
TITLE OF INVENTION: 511k Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Comparible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,747
FILING DATE: 14 MAR-1994
CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: MUTPHY Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELEPHONE: 703-205-8000
TELEPHONE: 703-205-8050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: N. clavipes
TISSUE TYPE: minor ampullate gland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 2:
SEGUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                          450 -GVGSGASAASAAA 462
| | | | | | | | | |
                                                                                                                566 QGAGQGAGAAAAA 579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1..309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22040-3487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HYPOTHETICAL: NO FRAGMENT TYPE: 1. ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
US-08-209-747-2
                                                                                                                                                                         RESULT 12
US-08-209-747-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       93
  à
                                    셤
                                                                              å
                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
```

```
158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                 25;
                                                                                                                                                                                                                                                                                                                                                                                                       248 -----GAGAAAAAGGA-GQGGYG------GLGSQGAGRGGQ------GAGA 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGLGGQGVGAGGLGGQ-----GAGAAAAVGAGQGGYGGVGSGA----SAASAAASRL 465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----GRGAGAGGYGGQGGYGAGAGAGAAAAAGAGAGGYGDKEIACWSRCRYTVASTTSRL 738
                                                                                                                                                                                                                                                                                                                                                                ------RGGLGGQGAGAAAAAAAAGGAGQGGLGGQ-----GAGAAAAAAGGAGQG 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLGGQGAGQGAGAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAAGGAGQGGT 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     628 GAGGYGRGAGAGAAAGAGAAAGAGAGGGGGGGGGGA----GAGAAAAAGAGAGRGGY 682
                                                                                                                                                                                                                                                                                                                                         LGSQGAGRGGQ-----GAGAAAAAGGAGQGGYGGLG------SQGAGRGGLGGQ--
                                                                                                                                                                                                                                                                                                                                                                                         ----GAGAAAAAGGVGQGGLGGQGAGQGAGAAAAAGGAGQGGYGGLG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----SQBAGRGGSG--GQGAGAAAAAGGAGQGGYGGLG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            193 GAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGLGGQGAGAAAAG---GAGQGGLGGQ--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             282 AAAAAGGAGGGYG--------GOGAGGGGYGGLGSQGAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SSPQASSRVSSAVSNLVASGPINSAALSSTISNVVSQIGASNPGLSGCDVLIQALL 521
                                                                                                                                                                                                                                       Length 832;
                                                                                                                                                                                                                                                                                         2 ASMTGGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAAGGAGGGGYG-----
                                                                                                                                                                                                                                                                143; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----SQGAGRGGLGGQ-----GAGAAAAAAAG------
                                                                                                                                                                                                                                       DB 1;
                                                                                                                                                                                                                                      49.8%; Score 1343.5; DB llarity 47.9%; Pred. No. 9.4e-92, Conservative 19; Mismatches 14:
                                                                                                                                                ORGANISM: N. clavipes
TISSUE TYPE: minor ampullate gland
                                                LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: single
 703-205-8000
TELEPHONE: 703-205-800
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
                                                                                                                                                                                        CDS
1..309
                                                                                     linear
                                                                                                                                                                                                                                                   Similarity
                                                                                                                                      ORIGINAL SOURCE:
                                                                                                                                                                                    NAME/KEY:
LOCATION:
                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                                   Best Local Sim
Matches 343;
                                                                                                                                                                                                              US-08-458-298-2
                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                          FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                          137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        739
                                                                                                                                                                                                                                                                                                                                                                요
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

RESULT 14

```
30;
                                                                           APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Xu, Ming
APPLICANT: Xt, Ming
APPLICANT: STATE
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGAGQ---GGYG----GLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGGQGAGQG 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----GGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLG-GQ----GAGAAAAGGV---- 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63 PGQQGPSGPGSAAAAAASGQQGPGGYGPRQQGPGGYGQGQQGPSGPGSAAAASAAASAE 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          123 SGQQGPGGYGPGO------QGPGGYGPGQGPGGYGPGQGPSGPSGPGAAAAAAA 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    232 GY-----GPGQQGLSGPGSAAAAAGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAGP 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGYG----GLGSQGAG-RGGQGAG-AAAAAAGGAGQGGYGGQGAGQGGYGGLGSQGAGRG 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----AAAAAG----GAGQGG 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    104 -GQGGLGGQGAGQGAGAAAAAGGAGQGGYG----GLGSQGAGRGGSGGQG-AGAAAAA 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               211 GYGGLGSQGAGRGGLGGQGAGAAAAGGAGQGGLGGQGAG------AAAAAAGGAGQ 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17 GYG----GLGGQGAGQ----GAG-AAAAAAGGAGQGGYG------GLGSQGAGR-- 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 GYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAAAGPGGYGPGGYGPGGYGPGQQGPGRYG 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 42.4%; Score 1144; DB 1; Length 595; Best Local Similarity 51.1%; Pred. No. 3.4e-77; Matches 313; Conservative 25; Mismatches 139; Indels 136;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                   ADDRESSEE: Birch, Stewart, Kolasch & Birch STREET: 301 No. 5728810th Washington Street CITY: Falls Church
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLGGQGAGAAAAAAGGAGGGLGGQGAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US/08/425,069
FILING DATE: 19-APR-1995
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
Sequence 4, Application US/08425069
Patent No. 5728810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 14
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPHONE: (703) 205-8000 TELEFAX: (703) 205-8050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO: 4: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGIH: 595 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; MOLECULE TYPE: protein US-08-425-069-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                                                                                                             Virginia
                                                         GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                         U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                          STATE: V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
```

```
262 GGYG----GLGSQGAG-RGGQGAG-AAAAAAGGAGQGGYGGQGAGQGGYGGLGSQCAGRG 315
                                                                                                                                                                                                                                                                63 PGQQGPSGPGSAAAAAAGSGQQGPGGYGPRQQGPGGYGQGQQGPSGPGSAAAASAAASAE 122
                                       104 -GOGGLGGQGAGQAGAAAAAGGAGQGGYG----GLGSQGAGRGGSGGQG-AGAAAAA 157
                                                                                                                                         172 SGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAASGPGQQGPGGYGPGQQGPG 231
                                                                                                                                                                                                                                                                                                                  316 GLGGQGAGAAAAAAAAGGAGQGGLGGQGAG-----AAAAAAAG----GAGQGG 358
                                                                                                                                                                                                                                                                                                                                                                                   359 LGGQGAGQ----GAGAAAAAAAAAGGVRQGGYGGLGSQGAG-----RGGQGAG-AAAA 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAGGAGGGGTG----GLGGQGV-----GAGGLG------GQGAGAAAAVGAGQGG 447
                                                                                                                                                                          211 GYGGLGSQGAGRGGLGGQGAGAAAAGGAGGGGLGGQGAG------AAAAAAGGAGQ
                                                                                                       158 GGAGQ---GGYG----GLGSQGAGRGGLGGQCAGAAAAAAAGGAGQGGYGGLGGQGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search completed: December 18, 2002, 16:23:13 Job time : 16:1566 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PGLSGCDVLIQAL 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           508
                                                                       g
                                                                                                                                     g
                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                       å
                                                                                                     ò
                                                                                                                                                                          å
                                                                                                                                                                                                     g
                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30;
                                                                                                                                                                                                                                                                                                                                                          Sequence 4, Application US/08317844B
Patent No. 598984
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: ALIMON, Ming
APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
                                                                                                                                                                     343 GPGGYGPGSASAAAAAAGPGQQGPGGYGPGQQGPSGPGSASAAAAAAAAAAAGPGGYGPGQQG 402
                                                   407
                                                                                                   408 AAGGAGGGTG----GLGGQGV-----GAGGLG------GQGAGAAAAVGAGQGG 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----GGGAGAAAAAAGGAGGGYGGLGSQGAGRGGLG-GQ----GAGAAAAAGGV---- 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17 GYG----GLGGGGAGQ----GAG-AAAAAAGGAGGGGYG-------GLGSGGAGR-- 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 GYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAAGPGGYGPGQQGPGGYGPGQQGPGRYG 62
                                 359 LGGQGAGQ----GAGAAAAAAAAAAGGVRQGGYGGLGSQGAG-----RGGQGAG-AAAA
                                                                                                                                   460 AAASAGPGGYGPAQQGPSGPGIAASAASAGPGGYGPAQQGPAGYGPGSAVAASAGAGSAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels 136;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 595;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,844B
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 42.4%; Score 1144; DB 2; Best Local Similarity 51.1%; Pred, No. 3.4e-77; Matches 313; Conservative 25; Mismatches 139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFENCE/DOCKET NUMBER: 1447-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 241-1300
MURICALION (703) 241-2848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEX: 248345
INFORMATION FOR SED ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                         508 PGLSGCDVLIQAL 520
                                                                                                                                                                                                                                                          111111111111
576 PGLSGCDVLIQAL 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                              US-08-317-844B-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-317-844B-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           57
a
                                                                                                                                 g
                                                                                                                                                                                                 g
                                 ò
                                                                g
                                                                                                   ò
                                                                                                                                                                     ò
                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
```

OM protein - protein search, using sw model

December 18, 2002, 16:21:10 ; Search time 7.29394 Seconds Run on:

(without alignments) 1213.493 Million cell updates/sec

Title: Perfect score:

US-09-490-291-8 2700 1 MASWIGGQQMGRIRIRGYGG......GLSGCDVLIQALLGHHHHH 528 Sequence:

Scoring table:

106657 seqs, 16763532 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

Total number of hits satisfying chosen parameters:

length: 0 length: 2000000000 Minimum DB seq Maximum DB seq Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

Published_Applications_AA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description	Sequence 1, Appli		Sequence 6, Appli	Sequence 4, Appli	Sequence 10, Appl	4	Sequence 35807, A	_	Sequence 2, Appli		Sequence 6, Appli		Sequence 36720, A	Sequence 1, Appli	Sequence 22, Appl	Sequence 14, Appl	18,	Sequence 19, Appl	18,
ΩI	US-09-861-597-1	US-09-861-597-8	US-09-861-597-6	US-09-861-597-4	US-09-861-597-10	US-09-861-597-2	US-09-864-761-35807	US-10-052-586-97	US-09-804-898-2	US-10-029-217A-24	US-09-812-382-6	US-09-756-071B-20	US-09-864-761-36720	US-10-096-961-1	US-09-848-990-22	US-09-760-364-14	US-09-798-584-18	US-09-967-624-19	US-09-998-667-18
DB	10	10	10	10	10	10	10	12	10	σ	10	10	10	σ	10	10	10	10	10
% Query Match Length DB	651	909	909	909	714	529	357	1894	4679	3907	450	720	283	2211	201	201	200	200	200
% Query Match	74.0	8.69	69.4	68.8						19.8	19.6	19.4	19.4	18.9	18.5	18.5	18.5	18.5	18.5
Score	1999	1883.5	1873.5	1858.5	970	915.5	619	649	557	534	529	524	523.5	511	500.5	500.5	499.5	499.5	499.5
Result No.	-	7	m	4	5	9	7	80	6	10	11	12	13	14	15	16	17	18	19

App,	Appl	Appli	Appl	Appl	Appl	Appl	App11	App11	App11	Appl	Appl	Appl	Appl	Appl	Appl	525, App	Appl	Appl	App ,	Appl	Appl	Appl	App1	Appl	App1.
269,	37,	~	37,	37,	34,	34,	7	m	'n	40,	40,	18,	18,	35,	33,	525	39,	39,	451	38,	38,	25,	36,	25,	36,
Sequence	Sequence			Sequence						Sequence	Sequence	Sequence	Sequence	Sequence										Sequence	Sequence
186-269	157-37	154B-2	169A-37	21A-37	9-837-969A-34	21A-34	97-7	97-3	97-5	69A-40	21A-40	69A-18	21A-18	69A-35	21A-35	86-525	69A-39	21A-39	86-451	69A-38	21A-38	69A-25	69A-36	21A-25	21A-36
US-10-052-586	Ō	3-090-60-sn	0S-09-837-9	US-09-841-3	US-09-837-5	US-09-841-3	US-09-861-5	US-09-861-5	US-09-861-5	US-09-837-9	US-09-841-3	US-09-837-5	US-09-841-3	US-09-837-9	US-09-841-3	US-10-052-586-525	US-09-837-969A-39	US-09-841-3	US-10-052-5	US-09-837-5	US-09-841-3	US-09-837-5	US-09-837-9	US-09-841-3	US-09-841-3
12	10	10	10	10	10		10					10				12					10	10	10	10	10
1300	1400	1497	782	782	2003	2003	101	101	101	605	605	1255	1255	1085	1085	4440	1085	1085	1743	745	745	635	635	635	635
18.4	18.3	18.1	17.9	17.9	17.9	17.9	17.8	17.5	17.5	17.4	17.4	17.4	17.4	17.4	17.4	17.4	17.3	17.3	17.2	17.2	17.2	16.7	16.7	16.7	16.7
498	493	490	483	483	483	483	481.5	473	472.5	470.5	470.5	470.5	470.5	468.5	468.5	468.5	467.5	467.5	465.5	465	465	452	452	452	452
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	36	40	41	42	43	44	45

ALIGNMENTS

```
APPLICANT: GARSON, Jean-Claude
APPLICANT: GARSON, Jean-Pierre
APPLICANT: ARRAUDEAU, Jean-Pierre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
TITLE OF INVENTION: ANALOG
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/861,597
CURRENT APPLICATION NUMBER: 09/247,806
PRIOR FILING DATE: 1999-02-11
PRIOR PILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 1
LENGTH: 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66 AAAGGAGQGGYGGLGSQGAGRGCLGGQGAG--AAAAAAGGYGQGGLGG---QGAGQGA-GA 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 188; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6 GGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGCQGGAGAAA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 651;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        74.0%; Score 1999; DB 10;
68.2%; Pred. No. 8.2e-113;
tive 5; Mismatches 10;
Sequence 1, Application US/09661597; Patent No. US/200064339A1
GENERAL INFORMATION:
APPLICANT: PHILLIPPE, Michel
                                                                                                                                                                                                                                                                                                                                                                                                                                             ; ORGANISM: Nephila clavipes US-09-861-597-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 68.2 Matches 435; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

ò	167 GLGSQGAGRGGLGGQGAGA185	Db 140 GGLGSQGAGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGY 196
Q D	199 GLGSQGAGRGGLGGQGAGAAAAAAGGAGQGGLGGQGAGAGAAAAAAAAA	Qy 159 GAGQGGYGGLGSQGAGRGGLGGQGAGA 185
δλ		Db 197 GGLGSQGAGQGGYGGLGSQGAGGGGGGGGGGAAAAAAAAAGGGGGGGGGGGGG
ор	259 GSQCAGRGEEGAGAAAAAGGAGQGGYGGLGGQGAAQGGYGGLGSQGAGRGGLGGQGAG- 317	Qy 186
ò	233 AAAAGGAGGGLGGOGAGAAAAAAGAAGGGYGGLGSQGAGRGGOGAGA-AAA 284	Db 257 AAAGGAGGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGQGGYGGLGSQ 316
අ දි	318 AAAAGGGGGGGGGGGGGGGGGAGAAAAGGAGGGGGGGG	219 GAGRGGLGGGGGGADADAAGGGGGGGGGGAGAAAAAGGGGGGGG
<u> </u>	AAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	
λŏ	AA AAAAAGGAGO	
qq	 438 Aaaaaaggagggygglcnogagrggggaaaaaggaggggggggggggggggggagaaa 497	Qy 330 AAGGAGGGLGGQGAGAAAAAAGGGGGGGGGQGAGQGAGAAAAAAAA
δ i	337	436 AAGGAGGGEGSQGAGAGAAAAAAGGAGGGGGGGGGGGGGG
8 8	 * 7	Oy 383 ROGSYGGLGSOGARRGGG 401 11 1 1 1 1 1 1
g		
à á	GAGGLGGGGAGAAAAVGAGGGGYGGYGSGASAASAAAS 	Db 555 AGAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGG
8	614 GRGGLGGQGAGAAAGGAGQGGYGGVGSGASAASAAA 651	0000111 2
RESU US-0 ; Sec ; Pa	RESULT 2 105-09-861-597-8 5 Sequence 8, Application US/09861597 5 Patent No. US20020064539Al 5 GENERAL INFORMATION: 7 APPLICANT: PHILLIPPE, Michel	US-09-861-597-6 Sequence 6, Application US/09861597 Sequence 6, Application US/09861597 Sequence 6, Application US/09861597 Seneral Incommarion: APPLICANT: PHILLIPPE, Michel APPLICANT: GARSON, Jean-Claude APPLICANT: ARRAUDEAU, Jean-Pierre
		; TILLE OF INVENTION: COSMETT ON DERMATOLOGICAL COMPOSITION CONTACTING AT ; TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN ; TITLE OF INVENTION: AND STATEMENT OF THE REPRESENCE: 6348-0345-0
	TITLE OF INVENTION: ANALOG FILE REFERENCE: 6388-0365-0 CHIRENT APPLICATION NUMBER: US/09/861,597	; CURRENT APPLICATION NUMBER: US/09/861,597 ; CURRENT FILING DATE: 2001-05-22 ; PRIOR APPLICATION NUMBER: 09/247,806
	PRIOR APPLICATION NUMBER: 09/247,806 PRIOR FILING DATE: 1999-02-11 PRIOR APPLICATION NUMBER: FR 98/01614	; PRIOR FILING DATE: 1999-02-11 ; PRIOR APPLICATION UMBER: FR 98/01614 ; PRIOR FILING DATE: 1998-02-11 : NUMBER OF SEO ID NOS: 14
	PRIOR FILING DATE: 1998-02-11 NUMBER OF SEQ ID NOS: 14 SOFTWARE: Patentin Ver. 2.1	; SOFTWARE: Patentin Ver. 2.1 ; SEQ ID NO 6 ; LENGTH: 606
S		; TYPE: PRT ; ORGANISM: Nephila clavipes US-09-861-597-6
, us-0;	lavipes	atch 69.4%;
Ou Beg Mat	Query Match 69.8%; Score 1883.5; DB 10; Length 606; Best Local Similarity 69.6%; Pred. No. 5.9e.106; Matches 412; Conservative 3; Mismatches 26; Indels 151; Gaps 16:	Matches 410; Conservative 3; Mismatches 27; Indels 151; Gaps Ov 6 GGOOMGRIBTREVEGIGGOGAGABABABAGGAGGGGGGGGGGGGGGGGGGGGGG
٥y	RGYGGLGGQGAGQAAAAAAAGGAGQGGYGGLGSQGAGRGGQG 60	23 GCQGAGAAAAAAGGAGGGGGGGGGGAAAAAGGAGGGGGGGG
q ₀	GGQGAGAAAAAAAGGGGGLGSQGAGQGAG-AAAAAAGGAGQGGYGGLGSQGAGRGGQG	Oy 61 AGAAAAAGGAGGGGYGGLGSQGAGRGGLGGGGGGAGAAAAAGGVGQ 105
o d	61 AGAAAAAGGACQGGYGGLGSQGAGRGGLGGGGAGAAAAAGGGQ 105 	Db 82 AGAAAAAAAAAAGAGGGGGGGGGGGGGGGGGGGGGG
Οy		

16;

218 AGGACGCGGGGGGGGGGGGAAAAAAAGGGGGGGGGGGG	446 394 506 450 566	WESULT'S PARSULT'S RESULT'S RESULT'S RESULT'S RESULT'S RESULT'S RESULT'S RESULT'S Sequence 10, Application US/09861597 Sequence 10, Application US/09861597 Sequence 10, Application US/09861597 GENERAL INFORMATION: APPLICANT: BHILLIPPE, Michel APPLICANT: GARSON, Jean-Claude APPLICANT: GARSON, Jean-Claude TITLE OF INVENTION: COSMETIC ON DEFENANCE OR RECOMBINANT SPIDER SILK OR AN TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN INTERESTERENCE: 6388-0365-0 CURRENT APPLICATION NUMBER: US/09/861,597 CURRENT APPLICATION NUMBER: P8/01614 PRIOR PRIOR APPLICATION NUMBER: FR 98/01614 PRIOR PRIOR APPLICATION NUMBER: FR 98/01614 SEQ ID NO 10 SEQ ID NO 10 LENGTH: 714 TYPE: PRT CORGANISM: ALTIFICIAL Sequence FEATURE: CORGANISM: ALTIFICIAL Sequence FEATURE: CORGANISM: ALTIFICIAL APPLICATION OF ALTIFICIAL SEQUENCE: FEATURE: CORGANISM: ALTIFICIAL APPLICATION OF ALTIFICIAL SEQUENCE: FEATURE: CORGANISM: ALTIFICIAL SEQUENCE	Query Match 35.9%; Score 970; DB 10; Length 714; Best Local Similarity 45.0%; Pred. No. 1.8e-51; Matches 285; Conservative 18; Mismatches 138; Indels 192; Gaps 31;	OY 17 GYGGLGSQCAGQGAGAAAAAAAGGAGGCYGGLGSQCAG 55 1 1 1 1 1 1 1 1 1 1	102 -GVGQGGLGGQGAGQGAG-AAAAAAGGAGGGGGGGLGSQGAGGGGGGGGLGSQGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	AAAGPGGYGPGQQGPGGYGPG -GLGSQGAGRGGLGGQCAGAA
159GAGQGGYGGLGSQGAGRGGLCGQGAGA	0y 330 AAGGAGGGELG GOGAGGAAAAAAAGGAGGGELGG GOGAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	RESULT 4 US-09-861-597-4 Sequence 4, Application US/09861597 Sequence 4, Application US/09861597 Batent No. US20020064539a1 GENERAL INFORMATION: APPLICANT: PHILLIPPE, Michel APPLICANT: ARRAUDEAU, Jean-Claude APPLICANT: ARRAUDEAU, Jean-Plerre TILLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT TILLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN TILLE OF INVENTION: ANALOG TITLE OF INVENTION: ANALOG TITLE OF INVENTION: ANALOG TITLE OF INVENTION NUMBER: US/09/861,597 CURRENT FILING DATE: 2001-05-22 PRIOR PPLICATION NUMBER: ER 98/01614 PRIOR PELING DATE: 1999-02-11 PRIOR PELING DATE: 1998-02-11 NUMBER OF SEQ ID NOS: 14 SOFTWARE: PATENTIN Ver. 2.1 SEQ ID NO 4 LENGTH: 606 TTVPE: PRT CORGANISM: Nephila clavipes US-09-861-597-4	Query Match 68.8%; Score 1858.5; DB 10; Length 606; Best Local Similarity 73.8%; Pred. No. 1.8e-104; Matches 409; Conservative 5; Mismatches 31; Indels 109; Gaps 18;	OY 6 GCOOMGRIRIRGYGGLGGGGGGGGGGGAAAAAAAGGGGGGGGG 48	107 GLGGQGAGGAAAAAGGAGGGGGGGGGGGGGGGGGGGGGG	Db 158 GLGSQGAGGABAAAAAGGAGGGGGGGGGGGGGGGGGGGGGGG

QOGPSGPGSAAAAAAGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAGPGGY 296

```
TILLE OF LINEARIENT GENE EXPRESSION ANALYSI:
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR PULICATION NUMBER: US 60/180,312
PRIOR PELING DATE: 2000-02-04
PRIOR PELING DATE: 2000-05-26
PRIOR FILING DATE: 2000-06-26
PRIOR PELING DATE: 2000-06-36
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-03
PRIOR PILING DATE: 2000-09-03
PRIOR PELING DATE: 2000-00-04
PRIOR PELING DATE: 2000-00-07
PRIOR APPLICATION NUMBER: PCT/US01/0066
PRIOR PILING DATE: 2001-01-30
PRIOR PILING DATE: 2001-01-30
PRIOR PELING DATE: 2001-01-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: PCT/US01/00661
FLING DATE: 2001-01-30
APPLICATION NUMBER: PCT/US01/00670
FILING DATE: 2001-01-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 35807, Application US/09864761
Patent No. US20020048763A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               450 GVGSGASAASA 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           519 GPGSQASAAAS 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -09-864-761-35807
                                                                                                                                                                                               264
                                                                                                                                                                                                                                                                                                                          318
                                                                                                                                232
                                                                                                                                                                                                                                                                                                                                                                                                                                                    361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR
                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
   셤
                                                                  ö
                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-861-597-2

Sequence 2, Application US/09861597

Sequence 2, Application US/09861597

Patent No. US2002006453941

GENERAL INFORMATION:

APPLICANT: GRSON, Jean-Flerre

APPLICANT: ARRADBAU, Jean-Flerre

TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT

TITLE OF INVENTION: ANALOG

FILE REFERENCE: 6388-0365-0

CURRENT APPLICATION NUMBER: US/09/861,597

CURRENT APPLICATION NUMBER: 09/247,806

PRIOR FILING DATE: 1999-02-11

PRIOR APPLICATION NUMBER: FR 98/01614

PRIOR FILING DATE: 1999-02-11

PRIOR FILING DATE: 1998-02-11

NUMBER OF SEQ ID NOS: 14

SEQ ID NO 2:

SEQ ID NO 3:

SEQ ID NO 3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29;
                                                                                                                       223 GGLGGQGAGAAAAAG------GAGQGGLGG------QGAGAAAAAAGGAGQGGYG 265
                                                                                                                                                                  ---RGGQGAGAAAAAAGGAGGGTGGL-----GGQGVGAGGLGGQGAGAAAAVGAGQG 446
                                              266 GLGSQGAGRGGQG-----AGAAAAAAGGAGQGGYGGQGAGQGGYGGLGSQGAGRGGLG 318
                                                                                                                                                                                                                                                                                            ----- 352
                                                                                                                                                                                                                                                                                                                                                                                                                             CAGQGGLGGQGAGQ----GAGAAAAAAAGGVRQGCXG------GLGSQGAG--- 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    535 GPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGG---PGGYGPGQQGPGGYGPGQQPGGYG 591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63 PGQQGPSGPGSAAAAAAGSGQQGPGGYGPRQQGPGGGGGGGGGPGSGPGSAAAASAAE 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      105 QGGLGGQGAGQGAGAAAAAAGGAGQGGYG----GLGSQGAGRGGSGGQG-AGAAAAAAGG 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SCCPGGYGPGQ-------QGPGGYGPGQQFPGGYGPGQQGPSGPGGAAAAAAAGG 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        57 ----GGQGAGAAAAAAGGAGQGGYGGLG--SQGAGRGGLGGQG-----AGAAAAAGGVG 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17 GYG----GLGGQGAGQ----GAG-AAAAAAAGGAGGGGYG-------GLGSQGAGR-- 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 GYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAAGPGGYGPGQQGPGRYG 62
193 GAGQ---GGYG----GLGSQGAGQ---GLGSQGAGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 131;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 529;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Score 915.5; DB 10;
; Pred. No. 2.4e-48;
20; Mismatches 136; I
                                                                                                                                                                                                                                                                                                                                                                                  319 GQG-AGAAAAAAAGGAGQGGLGGQGAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    652 GYGPGQGPGGYGPGQQGPSGPGSAAAAAAA 684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33.9%;
Similarity 47.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 47.9%
Matches 264; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; ORGANISM: Nephila clavipes US-09-861-597-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 123
                                                                                                                       à
                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
   à
                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
APPLICANT: Rank, David R.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
160 AGQ---GGYG----GLGSQGAGRGGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY
                                           231
                                                                                       263
                                                                                                                                   285
                                                                                                                                                                             317
                                                                                                                                                                                                                      342
                                                                                                                                                                                                                                                                  360
                                                                                                                                                                                                                                                                                                         343 GGYGPGSASAAAAAAGPGQQGPGGYGPGQQGPSGPGSASAAAAAAAAGPGGYGPGQQGPG 402
                                                                                                                                                                                                                                                                                                                                                     GQGAGQ----GAGAAAAAAAAAAGGVRQGGYGGLGSQGAG-----RGGQGAG-AAAAA 409
                                                                                                                                                                                                                                                                                                                                                                                                 103 GYAPGOOGPSGPGSASAAAAAAAAG--PGGYGP-GQQGPGGYAPGQQGPSGPGSAAAAA 459
                                                                                                                                                                                                                                                                                                                                                                                                                                          410 GGAGQGGTG----GLGGQGV-----GAGGLG------GQGAGAAAAVGAGQGGYG 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YG----GLGSQGAG-RGGQGAG-AAAAAAGGAGQGGYGGQGAGQGGYGGLGSQGAGRGGL
                                                                                                                                                                                                 172 PGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAASGPGQQGPGGYGPGQQGPGGY
                                                                                     213 GGLGSQGAGRGGLGGQGAGAAAAGGAGQGGLGGQGAG------AAAAAAGGAGQGG
```

```
19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97 AAAAGGVGQGGLGGQGAGAGAAAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAA 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S-cocodececerersenecececececesenage----esesenee-ceneeceneece 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             277 QGAGAAAAAAGGAGGGYGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAGGAGQ 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        37 AAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGA 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17 ASGGSVOCGGGGDGGDG---GGSGSGGDAGGGGGDGGGGDSGGGDGGGG----GGSGDG- 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 157 ACGACQCGYGGLGSQCAGRGCLGCQCACAAAAAAGGAGQCGYGGLGCQCACQCGYGCLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         217 SQGAGRGGLGGQGAGAAAAAGGAGQGGLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         397 RGGQGAGAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAGAAAAVGAGQGGYGGV 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SGGG-----GGSGDGGGGSGSGDGGGGGGRGFSGDGG----GGGGGGTGGV 341
                                                                                                                                                                                                                                               TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
COTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
OTHER INFORMATION: EXPRESSED IN FIGNAL = 34
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 28
OTHER INFORMATION: EXPRESSED IN PETAL SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 25
OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 25
OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 25
OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 21
US-09-864-761-35807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 25.1%; Score 679; DB 10; Length 357; Best Local Similarity 43.6%; Pred. No. 2.1e-34; Matches 181; Conservative 12; Mismatches 132; Indels 90
                     PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR PELICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33807
LENGTH: 357
PRIOR APPLICATION NUMBER: US 60/234,687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 97, Application US/10052586
Patent No. US20020127584A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Chen, Jian
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Audrey
APPLICANT: Garney, Austin L.
APPLICANT: Garney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Smith, Victoria
APPLICANT: Smith, Victoria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
US-10-052-586-97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

MBER: 60/08107 998-04-08 MBER: 60/08119 998-04-19 MBER: 60/08183 998-04-15 MBER: 60/08256 MBER: 60/08256 MBER: 60/08270 MBER: 60/08270 998-04-21 MBER: 60/08270	NUMBER: 60 11998-04-2 10998-04-2 10998-04-2 10998-04-2 10998-04-2 10998-04-2 10998-05-2	
PLICATION LING DATE LING DATE LING DATE LING DATE LING DATE PLICATION LING DATE LING DATE PLICATION LING DATE	PPLICATION ILING DATE ILING DATE PPLICATION ILING DATE	MATCH DATE MATCH M
NARARARARA NICOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCO	PRIOR	PRIOR

946 AGGIGATGCACCAGATTTAGTTGATGAGAGAAGATGAAAGTGCAGAGCATGATGAATA 1005 222 ---RGGLGGQGAGAAAAGGAGQGGLGGQGAG---------AAAAAAGGA 259 719 AGGGAAATTAAAAGGC-----TGAAAAAAGAGAAAACCAGAGGAGGAAGTAAAGAA 768 R-----GGLGGQGAGAAAA------GGVGQGGLGGQGAGQGA 117 769 ATTGAAACCCAAAAGGCACAAAAATTTTAGTTTACTTTCATTTGGAGGAAG---CTGA 825 886 TCATGACTTGCTTAAGGATGATCCACATCTCCAGTTCTGTTCCAGTTGTAGAAAGTGAAAA 945 26 AGQGAGAAAAAAAAGAAGQGGYGGLGSQGAGRGGQGAGAAAAAAAGGAGQGGYGGLGSQGAG 85 118 GAAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGA-----AAAAAGGAGQGGYGG 168 LGSQG-----AGRGGLGGQGAGAAAAA Indels 200; Length 1894; Query Match 24.0%; Score 649; DB 12; 1 Best Local Similarity 32.8%; Pred. No. 5.4e-32; Matches 216; Conservative 12; Mismatches 230; PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR FILING DATE: 1998-06-04
PRIOR FILING DATE: 1998-06-04
PRIOR PILING DATE: 1998-06-04
PRIOR PILING DATE: 1998-06-07
PRIOR PAPLICATION NUMBER: 60/08873
PRIOR PILING DATE: 1998-06-10
PRIOR PILING DATE: 1998-06-10
PRIOR PILING DATE: 1998-06-10
PRIOR PILING DATE: 1998-06-10
PRIOR PAPLICATION NUMBER: 60/08811
PRIOR PILING DATE: 1998-06-10
PRIOR PAPLICATION NUMBER: 60/08824
PRIOR PILING DATE: 1998-06-10
PRIOR PAPLICATION NUMBER: 60/08863
PRIOR PAPLICATION NUMBER: 60/08864
PRIOR PILING DATE: 1998-06-11
PRIOR PILING DATE: 1998-06-11
PRIOR PILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089165
PRIOR PILING DATE: 1998-06-12
PRIOR PILING DATE: 1998-06-16
PRIOR PILING DATE: 1998-06-17
PRIOR PILING DATE: 1998-06-17 FILING DATE: 1998-06-05 98 g ò g ò qq ò q ò g ò qq

19;

------GQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGG 293

us-09-490-291-8.rapb

GUGGGGARGEGGGGAGAAAAAAAGGGGGGGAAATTAAAGGGGGGAAATTCAGCGGGGAAAAGGGGGGGAGAAAGGGGGGGAGAAAGGGGGG	19.84 19.8	CURRENT APPLICATION NUMBER: 1 CURRENT FILING DATE: 2002-0. PRIOR FILING DATE: 2000-12-2. NUMBER OF SEQ ID NOS: 32 SOFTWARE: Patentin Ver. 2.1 SEQ ID NO 24 IENGTH: 3907 TYPE: PRT CRGANISM: MUS musculus	QY 433	
	73OGGYGGLGSQGAGRAAAAAGGVGQGGLGGQGAG	-09-804-898-2 Duery Match 20.6%; Score 557; DB 10; Length 4679; Best Local Similarity 28.8%; Pred. No. 3.5e-26; Matches 237; Conservative 15; Mismatches 229; Indels 342; Gaps 19 GGLGGQGAGQGAAAAAAAGGAGGGGGGGGGGGGGG	Sequence 2, Application US/09804898 Patent No. US20020045264A1 Saguence 2, Application US/09804898 Patent No. US20020045264A1 SAPLICANT: DURING, MATTHEW APPLICANT: DIRING, MATTHEW APPLICANT: DIRING, MATTHEW TITLE OF INVENTION: PRODUCTION OF CHIMERIC FILE REFERENCE: 102182-14 CURRENT APPLICATION NUMBER: US/09/804,898 PRIOR FILING DATE: 2001-03-13 PRIOR FILING DATE: 2001-03-14 NUMBER OF SEQ ID NOS: 6 SOFTWARE: Patentin version 3.0 SEQ ID NO 2 LENGTH: 4679 TYPE: PRT ORGANISM: adeno-associated virus 2	1066 CACAAGTGCGAATGTTAAATCAGTGGAGAACAAAAAAGGAGGGG

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AND COMPOSITIONS RELATING TO A CARDIAC-SPECIFIC R REGULATORY FACTOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CGCCACCTCTATCCTGCAAAGGCTGGCGAGGTG----GT 1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AG-GAGCTGAAGTTGCGATCACTGCCTGTCTCGGGCACCA 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----AGAAAAAGGAGQGGYGGLGSQGAGR 142
-----GLGGQGAG 323
                                                                       CTACTTCCACATGCACGTGCTCGTGGAAACCACCGGGGTG 620
                                                                                                                                                                                CTTTGCCAAACTGGTTCGCGGTCACAAAGACCAGAAATGG 739
                                                                                                                                                                                                                        | | | | | | : |:| | | | : : : CTCCAGTGGGGCGTGGACTAATAGGAAC-----AGTATT 846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GRGGLGGQGAG------AAAAAGGVGQGGLGGQG 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGYGGLGSQGAGRGG-----LGGQG--AGAAAAAAGG 193
                                                                                                                                                                  AAAAVGAGQG-GYGGVGSGASAASAASRLSSPQASSRVS 475

    8; Score 534; DB 9; Length 3907;
    8; Pred. No. 7e-25;
    4; Mismatches 212; Indels 226; Gaps

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGAGQGGYGGLGSQGAGR-------GGQGAGAA 64
                                                                                                                                                                                                                                                                                                                                             I----SNVVSQIGASNPGLSGC 513
                                                                                                                                                                                                                                                                                                                                                                   ACGGAGCGTAAACGGTTGGTGGC 889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :: US/10/029,217A
:-03-19
60/257,761
                                                                                                                                                                                                                                                                                                                                                                                                                                                     /10029217A
```

Thu Dec 19 09:28:31 2002

```
356 QGGLGGQGAGQGAGAAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQG 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19 GGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAAGGAGQGGYGG 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12 GGAGGTTGGCCAGTCAATA----GGTTACTTTATGAGTTGCTAACCCTGGTGAGCAGG 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 524; DB 10; Length 720;
Pred. No. 6.6e-25;
4; Mismatches 204; Indels 272;
                                         ZIP: 44114

COMPUTER READABLE FORM:
MEDIUW TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                  APPLICANT: Tryggvason, Karl
Kallunki, Pekka
Pyke, Charles
TITLE OF INVENTION: Laminin Chains: Diagnostic Uses
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --0GAG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----LGSQGAGRGGSGGQGAGAAAAAGGAGQGG------
                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: Fay Sharpe Fagan Minnich & McKee STREET: 1100 Superior Ave, Suite 700 CITY: Cleveland
                                                                                                                 415 GVPGVGVPGVPGAGVPGAGVP---GGGVPGGGVPGGG 450
                                                                              116 GTGGLGGQGVGAGGLGGQGAGAAANGAGQGGYGGVGSG 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: US/09/756,071B
FILING DEPE: 08-Jan-2001
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REGISTRATION NUMBER: 24,175
REFERENCE/DOCKET NUMBER: TRV 20014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US 09/663,147
FILING DATE: 150-September 2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; TOPOLOGY: linear
; MOLECULE TYPE: proteln
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-756-071B-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----YGGLGSQGAGR--GGLGG---
                                                                                                                                                                                                           Sequence 20, Application US/09756071B Patent No. US20020052307A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: Minnich, Richard, J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: 216-861-5582
TELEFAX: 216-241-1666
INFORMATION FOR SEQ ID NO: 20:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 30.4%;
Matches 210; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE: Ohio
                                                                                                                                                                          RESULT 12
US-09-756-071B-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      165
                                         а
                                                                              ò
                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16;
                                                                                                                                                                                        1470 TGACCCTGCAGGCCTCGCCACTGCAGATCCTCGTGAAGGAGGAGGGCCCCCGGGCCGGGT 1529
                                                                                                                                                                                                                                                                                                                               1410 ATGAAAACTCCACCCCGGGGACACCTTTGGTGAGATGGTGACATCACCTCTGACGCAGC 1469
                                                                                                                                                                                                                                                                1350 CGGGCTCCACGCCCCCGTGTCTCCCACCCCTCGGAGCGCTCACTGCTCAGCACGGGCG 1409
                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                     QGAGQGAGAAAAAAAAAAAGGVRQGGYGGLGSQGA----GRGGQGAGAAAAAAGGAGQGG 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         76 YG--GLGSQGAGRGGLGGQGAGAAAAGGVGQGGLGGQGAGGAGAAAAAAGGAGQGGYG 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              134 --GLGSQGAGRGGSGGQGAGAAAAAAGGAGQGGYG--GLGSQGAGRGGLGGQGAGAAAA 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         190 AAGGAGQGG---YGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGA---GAAAAAGGAGQGG 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LGG----QGAGAAAAAAGGAGQGGYG--CLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGQG 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              299 AGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAAGGAGGGGGG---QGAGAAAAAAAAGGAG 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        17 GYGGLGGQGAG-QGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGG 75
   ------GAGGGAGAAAAAG--GAGGGGYGGLG
                                                                                                                                                    ------AGRGGYGGLGSQG-----AGRGGLGG--
                                                                                                                                                                                                                               -----QGAGAAAAAAAGGAGQG-----GLGGQGAGAAAAAGGA-GQGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 6, Application US/09812382
Patent No. US20010034050A1
GENERAL INFORMATION:
APPLICANT: Chilkcti, Ashutosh
TITLE OF INVENTION: Fusion Peptides Isolatable by Phase Transition
FILE REFERENCE: 4176-101
CURRENT APPLICATION NUMBER: US/09/812,382
CURRENT FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: US 60/190,659
PRIOR FILING DATE: 2000-03-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19.6%; Score 529; DB 10; Length 45(
35.9%; Pred. No. 2.2e-25;
.tve 16; Mismatches 234; Indels
                                                                                                                                                                                                                                                                                                                                                                                                     417 -TGGLGGQGVG------AGGLGGQG-AGAAAVGAGQG 446
                                                                            -GGQGAGAAAAAGGAGQGGYGGQGA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARE: Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 35.9%
Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT
; ORGANISM: ELP[V5A2G3-90]
US-09-812-382-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQ ID NOS:
                                                                          SQGAGR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -09-812-382-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQ ID NO 6
                                                                          269
                                                                                                                                                    300
                                                                                                                                                                                                                             320
                                                                                                                                                                                                                                                                                                     362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           246
                                     g
                                                                                                           g
                                                                                                                                                                                        a
                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                        g
à
                                                                          ö
                                                                                                                                                    ð
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

29;

Gaps

---A 185

g õ g ç ద ò g ò 엄 ò 셤 à g ò g å q

```
GENERAL INFORMATION:
APPLICANT: GUEGLER, KARI et al.
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
TITLE OF INVENTION: UNCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: CLO00849DIV
CURRENT APPLICATION UNBERS (25/10/096,961
CURRENT FILING DATE: 2002-03-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      224 GLGGQGAGAAAAAGGAGQGGLGGQGAGAAAAAAGGAG---QGGYGGLGSQGAGRGGQGAG 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             193 ------GGGGSDGGGDGGGGGGSDGCGGGSDRGGG---GGSDGGGGSDGGGG 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              104 GOGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAGGAGQG 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            164 GYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRG 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        43 GDGGGGSDGGGGGGGGGGSDGGGGGGSDG---GGSGSDGGG---GGSSDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
19.4%; Score 523.5; DB 10; Length 283;
Best Local Similarity 39.9%; Pred. No. 3.2e-25;
Matches 139; Conservative 7; Mismatches 123; Indels 79;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       239 GDGGG--SHGGGGGGGGGGGGGGGGGGAD------HDGGYG 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3.8° 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N: EXPRESSED IN LUNG, SIGNAL = 17
N: EXPRESSED IN PLACENTA, SIGNAL = 11
N: EXPRESSED IN HEART, SIGNAL = 9
N: EXPRESSED IN HEART, SIGNAL = 12
N: EXPRESSED IN HEART, SIGNAL = 12
N: EXPRESSED IN FETAL LIVER, SIGNAL = 10
N: EXPRESSED IN FOUT LIVER, SIGNAL = 10
N: EXPRESSED IN BOULT LIVER, SIGNAL = 9.4
N: EXPRESSED IN BONE MARROW, SIGNAL = 10
N: EXPRESSED IN BONE MARROW, SIGNAL = 10
N: EXPRESSED IN BRAIN, SIGNAL = 12
                                                   PRIOR APPLICATION UNMBER: PCT/USO1/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/USO1/00670
PRIOR PILING DATE: 2001-01-30
PRIOR PILING DATE: 2001-01-30
PRIOR PILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR FILING DATE: 2000-06-30
PRIOR PILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers: 1.1
SEQ ID NO 36720
      APPLICATION NUMBER: PCT/US01/00662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MAP TO AC006547.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-10-096-961-1
; Sequence 1, Application US/10096961
; Patent No. US20020155572A1
                                      LING DATE: 2001-01-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURE:

OTHER INFORMATION: DAY
OTHER INFORMATION: ED;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NESULT 13

RESULT 13

RESULT 13

RESULT 13

RESULT 13

RESTAL PAPLICATION US/09864761

FRACE NO. US200200486531

FRACE NO. US200200486531

RESTAL PROMATION SALTON G.

APPLICANT: Penn. Sharron G.

APPLICANT: Hanel. David R.

APPLICANT: Hanel. David R.

APPLICANT: Hanel. David R.

APPLICANT: HANGEL DAVID R.

TITLE OF INVENTION: HUNAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ITLE REPERENCE: ACHOLGA-Y.

PRIOR FILING DAFE: 2000-05-29

PRIOR PLICANTON NUMBER: US 60/29/464,761

PRIOR PLICANTON NUMBER: US 60/29/466

PRIOR PLICANTON NUMBER: US 60/20/456

PRIOR PLICANTON NUMBER: US 60/20/456

PRIOR PLICANTON NUMBER: US 60/20/466

PRIOR FILING DAFE: 2000-09-27

PRIOR FILING DAFE: 2000-09-27

PRIOR FILING DAFE: 2000-09-27

PRIOR FILING DAFE: 2001-01-30

PRIOR FILING DAFE: 2001-01-30

PRIOR PLICANTON NUMBER: PCT/US01/00666

PRIOR FILING DAFE: 2001-01-30

PRIOR PLICANTON NUMBER: PCT/US01/00666

PRIOR PRIOR APPLICANTON NUMBER: PCT/US01/00666

PRIOR PRIOR APPLICANTON NUMBER: PCT/US01/00666

PRIOR PPING APPLICANTON NUMBER: PCT/US01/00666

PRIOR PRIOR PLICANTON NUMBER: PCT/US01/00666

PRIOR PPING APPLICANTON NUMBER: PCT/US01/00666

PRIOR PRIOR APPLICANTON NUMBER: PCT/US01/00666
                                                                                                                                                                                                                                                                                                                                                                                                                                        300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           238 ---GAGQGGLGG---QGAGAAAAA---- 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301 -----QGGYGGLGSQGAGR-----335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             375 AAAAAGGVRQGGYGGLGSQGAGRG----GQGAGAA--AAAAGGAGQGGTGGLGGQGVGAG 428
182 GATGTCCTTTAAGCCGGAGCTGATTCGGGCTGCTTGTTTTTTGTGAGTTAGCGCTCTTA 241
                                                                                                                                                                                                                                             302 TGGACCACCCCGGGAGAGAAGGAGGCTCCGGGGAATCTCGCACATTCCAGGCAAAGGCT 361
                                                                                                                                                                                                                                                                                                                                                                       362 CCCGGGCCGCAGCCTCTGTGCCACACCCTTGGCCCGGGCCAGGTGTGCGCCCTCCTCGCT 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      336 -QGG------GAAAAAAAAAAAAAAAAAAAAAAAAAA 374
                                                                                                                      242 AGATTGGGCCTCCCAGTTTGAGGAAGGGGCGGGCTGCTGTCTACCTCTGTGAATCTGCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                     GYGGLGSQGAGRGGQGAGAAAAAAGGAGQG--------GYG----GQGAG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      601 GGAAAAGGAAGGCACAGGGGGGCGCAGAGTGAGAACCACCAACCGAGGCGCGGGGAGCG
                                                                                                                                                                                   209 -------GGGYGGLGSQGAGRGG---LGGQGA-------GAAAAG--
                                                             186 AAAAAAGG-----AGQGGYGGLGGQGAG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------GLGGQG--AGAAAAVGAGQGG 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 661 ACCCTGCAGCGGAGACAGAGTGAGCGG 690
                                                                                                                                                                                                                                                                                                                                                                                                                                     263
```

13;

```
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                        32;
                                                                                                                                                                                                                                                                                                                             126 GTTCCAGCGGCGCATCCCGCGTGCGCCCATCGCACATCTTCTTCGTGCAGCACATCGA 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             156 AAGGAGQGG-----YGGLGSQGAGRGGLGGQCAGAAAAAAGGA-GQGGYGGLG-- 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            306 CACGGCTGAGGAAGCGCAGCGCTGGGCGCGTCTGACCAAGCTCCGCGGGGGCGTGGA 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     366 CGCCATGAGCCAGCGGAGCGGCTAGACCACTGGATCCACTCCTATCTGCACCGGGTGA 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           546 CGACCGTCTAGAGGGGGCTGAGATCGAGGAGTTCCTGCGGCGGCGGCTGCTGAAGCGGCCGGA 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGLGGQ--GAGQGAGAAAAAAG---GAGQG---GYGGLGSQGAGRGGSGG--QGAGAAAA 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   246 AGCGCGCTGCCTCACCATCGCCTTCAAGGGCCGCCGCAAGAACCTGGACCTGGCGGCGCG 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----GQGGYG- 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----GLGSQGAGRGGQGAGAAAAAGGAGQ------GGYGGQGAGQG----- 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              426 CTCCAACCAGGACAGCAAGATGAGCTTCAAGGAGTCAAGAGCCTGCTGAGAATGGTCAA 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   303 -GYGGLGSQGAGRG-----AAAAAA 330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CGTGGACATGAACGACATGTACGCCTACCTCCTCTTCAAGGAGTGTGACCACTCCAACAA 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----- 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGQGAGQGAGAAAAAAAAAAGGVRQGGYGGLGSQGA----GRGGQGAGAAAAAAGGAG-Q 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CTGGAGGAGATCTTCCATCAGTACTCGGGCGAGGACCGCGTGCTGAGTGCCCCTGAGCT 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------QGAG-AAAAVGAG-----QGGYGGVGSGASAAASRLSSPQASSR 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       726 CATTCAGACCTATGAGCTCAACGAGACAGCCAAGCAGCATGAGCTGATGACACTGGATGG 785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         204 ------GQGAGQG-GYGGLGSQGAGRG-GLGGQGAGAAAAAG------ 237
                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                     23 GQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQ 82
                                                                                                                                                                                                                             18.9%; Score 511; DB 9; Length 2211; 28.9%; Pred. No. 1e-23; Live 14; Mismatches 249; Indels 278;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            474 VSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPG-LSGC 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 786 CTTC----ATGATGTACCTGTTGTCGCCGGAGGGGACTGC 821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------GAGQG-GLGGQGAGAAAAAGGA-----
                                                                                                                                                                                                                                                                                                                                                                               83 G----AGRGGLG------GQGAGAAAAAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        331 AGG----AGQGGLGGQGAGAAAAAAGGAGQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION NUMBER: 60/232,632
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 09/738,884
PRIOR FILING DATE: 2000-12-18
NUMBER OF SEQ ID NOS: 5
SOFTYARRE: FastSEQ for Windows Version 4.0
ERGI ID NO 1
LENGTH: 2211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
US-09-848-990-22
; Sequence 22, Application US/09848990
; Patent No. US20020048572A1
                                                                                                                                                                                                                                                 Best Local Similarity 28.9%
Matches 220; Conservative
                                                                                                                                                                       ; ORGANISM: Homo sapien
US-10-096-961-1
                                                                                                                                                         TYPE: PRT
                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                       106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           266
                                                                                                                                                                                                                                                                                                                                          윱
                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
```

```
9
                                                                                                                                TITLE OF INVENTION: Treatment of Hypertriglyceridemia and Other Conditions TITLE OF INVENTION: Using LXR Modulators
FILE REFERENCE: 018781-00491003
CURRENT APPLICATION NUMBER: US/09/848,990
CURRENT FILING DATE: 2001-05-03
FRIOR APPLICATION NUMBER: US 60/201,601
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PATENTING DATE: 2000-05-03
NUMBER OF SEQ ID NOS: 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          128 GOGGYGGLGSQGAGRGGSGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGCLGGQGAGAAA 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         248 GAGAAAAAAGGAGOGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGQGGYGGL 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     308 GSQGAGRGGLGGQGAGAAAAAAAGGAGGGLGGQGAGAAAAAAAGGAGGGLGGQGAGQG 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OTHER INFORMATION: Description of Artificial Sequence: flexible linker
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DCCATION: (1)...(97)
OTHER INFORMATION: Gly at positions 1-97 may be present or absent
NAME/KEY: MOD_RES
LOCATION: (105)...(201)
OTHER INFORMATION: Gly at positions 105-201 may be present or absent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         188 AAAAGGAGGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAGGAGQGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 18.5%; Score 500.5; DB 10; Length 201; Best Local Similarity 49.2%; Pred. No. 5.6e-24; Matches 118; Conservative 0; Mismatches 83; Indels 39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Search completed: December 18, 2002, 16:39:26
Job time : 13.2939 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
ORGANISM: Artificial Sequence
                             APPLICANT: Shan, Bei
APPLICANT: Schultz, Joshua
APPLICANT: Tu, Hua
APPLICANT: Tu, Hua
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-848-990-22
```

```
GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2002, 16:13:49; Search time 15.6009 Seconds

(without alignments)
3253.588 Million cell updates/sec

Title:
DS-09-490-291-8
Perfect score: 2700
Sequence: 1 MASMTGCQOMGRIRIRGYGG......GLSGCDVLIQALLGHHHHH 528
Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5
Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_73:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*

Database :

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

	Description	major ampullate fi	spidroin 2, dragl1	fibroin - Chinese				hypothetical glyci		_	_		_								hypothetical glyci	O,	hypothetical glyci			_	hypothetical glyci			
SUMMARIES	QI	A36068	A44112	T31328	F70806	B70807	D70807	E70806	E70895	A70869	S01820	A70896	B70812	A70812	E70917	F70963	A70807	H70987	H70846	E70820	A70934	QQBE31	G70917	G70846	F70824	A70893	F70620	E70824	6	D70835
	DB	7	7	7	7	~	~	7	7	~	-	7	7	7	7	7	7	7	~	~	7	-	~	~	7	7	~	~	~	7
	Length	718	627	2639	1901	1079	1489	1381	767	1660	465	823	882	749	1329	778	714	914	1538	923	1306	641	741	484	801	667	1011	783	731	957
æ	Query	81.1	42.5	S	34.6	34.3	33.8	33.0	32.6	32.2	32.1	31.8		31.5	31.4	31.3	31.2	31.1	31.0	30.9	30.8	30.7	30.6	30.4	30.4	30.3	30.1	30.1	30.0	29.9
	Score	2190	1148	968.5	935	927	912.5	892	880	868.5	866.5	828	857.5	850	847	æ	43	839.5	836	833	æ	828.5	27	820.5	13	818	œ	812.5	810	0
	Result No.	-	7	m	4		9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53

hypothetical glyci glycin-rich protei glycin-rich protei hypothetical glyci hypothetical glyci	hypothetical glyci elastin precursor glycine-rich cell loricrin - mouse
E70835 A70900 A70900 A70900 F70868 H70816 A70989 H70589 H70589 A70931 A260923	G70545 EAMS KNMU A35628
00000000000	7777
837 408 603 604 606 584 618 6113 6313	594 860 338 481
229.04 229.04 277.77 277.77 277.33	27.1 27.0 26.5 25.8
796.5 793.5 782.5 781.5 764.5 747.5 747.5 739.5 739.5	732 729.5 714.5 697.5
33 33 33 33 33 33 33 34 35 36 36 37 37 37 37 37 37 37 37 37 37 37 37 37	4 4 4 4 5 4 4 4

ALIGNMENTS	RESULT 1 A36068 major ampullate fibroin protein - orb spider (Nephila clavipes) (fragment) C;Species: Nephila clavipes C;Date: 08-Mar-1991 #sequence_revision 13-Jan-1993 #text_change 09-Sep-1997 C;Accession: A36068 R;Xu, M.; Lewis, R.V. Proc. Natl. Acad. Sci. U.S.A. 87, 7120-7124, 1990 A;Title: Structure of a protein superfiber: spider dragline silk. A;Reference number: A36068; MUD:90384959; PMID:2402494 A;Accession: A36068 A;Status: preliminary A;Molecule type: DNA A;Cross-references: GB:M37137; NTD:G150711; PID:G150712	A;Note: the authors translated the codon GGT for residue 292 as Gln, GTA for residue Query Match Query Match Best Local Similarity 70.0%; Pred. No. 5e-112; Matches 477; Conservative 6; Mismatches 10; Indels 188; Gaps 15;	QY 6 GGQQMGRIRIRGYGGLGGQGAGQAGAAAAAGAGGGGYGGLGSQCAGRGGGCAGAAA 65 :	Qy 66 AAAGGACOGGYGGLGSQQAGRGGLGGQGAGAAAAAGGGGGGGGGGOGAGGGA-GA 119 	Qy 120 aaaaaggagggggggggggggggggggggggggggagaaaa	QY 167 GLGSQGAGRGGLGGQGAGA	Qy 186ANAAAAGGAGGGYGGLGGQGAGGGGYGGLGSQGAGRGLGGGGAGA 232 	Qy 233 AANAGGAGGGGGGGGGAGAAAAAGGAGGGGGGGGGG	QY 285 AAGGAGGGCGGGG0CAGGG	OY 325 AA 336 1
------------	--	---	--	---	--	------------------------------------	---	---	------------------------------------	------------------

Qy 337	
RESULT 2 A44112 Spidtoin 2, dragiline silk fibroin - orb spider (Nephila clavipes) (fragment) N:Alternate names: silk fibroin, dragiline C;Species: Nephila clavipes C;Species: Nephila clavipes C;Species: Nephila clavipes C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Mar-1998 R;Hinman, M.B.: Lewis, R.V. J. Biol. Cham. 267, 19324, 1992 A;Accession: A44112; MulD:92406876; PMID:1527052 A;Accession: A44112 A;Accession:	kmoth hyl (Chli ence_re ; Yukuh; a Libral zation of sanslatee
30;	Query Match 35.9%; Score 968.5; DB 2; Length 2639; Best Local Similarity 42.6%; Pred. No. 2.2e-45; 164; Gaps 24; Matches 271; Conservative 24; Mismatches 177; 164; Gaps 24; Qy 20 GLGGGCAGGGAGAAAAAAAGGAGGCGGGGGGGGGG
AAAAGGAGGGGYGGLGSOGAGR 56 AAAAAAAAGPGYGPGQGPGGYGPGQGPGRYG 62 GSQGAGRGGLG-GQGAGAAAAGGV 103	
104 -GQGGLGGQGAGAAAAAGGAGGGGGGGLGSQGAGRGGSGGG-AGAAAAAA 104 -GQGGLGGGAGAAAAAAGAGGAGGGGGGLGSQGAGRGGSGGG-AGAAAAAA 123 SGQQGPGGXGPGQQGPGGYGPGQGFGYGFGQGGPGGPGSGPGSAAAAAAA 158 GGAGGLGSQGAGRGGLGGQGAAAAAAAGGAGGGYGGLGGGAGGG 172 SGPGQGFGGYGPGGYGFGGGGGGGGGGGGGGGGGGGGGGGG	
Qy 211 GYGGLGSQGAGRGAGAAAAGAGGGGLGGGGAG 1	Db 459 AAAAAAAAGGAGGGYGGYGGYGGYGGAAAAAAAAAGAGAGAG-GSYGWGDGGYG 516 Qy 279 A-GAAAAAAGGAGGGYGGGAGGGGGGGGGGGGGGGGGGG

362 QGAGQGAGAAAAAAAAAGGYR-QGGYGGLGSQGAGRGGQGACAAAAAAGGAGOGGTGGL 1:	Qy 351
Db 637 YGSDSAAAAAAAAAAAAAAAGGGGGGGGGGGSAAAAAAAAA	OY 387 YGGLGSOGAGRGGQGAGAAAAAGGAGQGGTGGLGGQGV 425
474 VSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPG 509 11 : : : : :	426
RESULT 4	OY 461 AASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCD 514
ttical glycine-rich protein Rv3508 - Mycobacterium tuberculosis (strain H37RV) les: Mycobacterium tuberculosis 17-Jul-1998 #sequence_revision 17-Jul-1998 #sequence_revision 17-Jul-1998 #sequence_revision 17-Jul-1998 #seq	RESULT 5 B70807
R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Bers, J.; Rutter, S.; Segger, R.; Skelton, S.; Squares, S.	hypothetical glycine-rich protein Rv3512 - Mycobacterium tuberculosis (strain H37RV) C;Species: Mycobacterium tuberculosis . C;Date: 17.5U1-1998 #sequence_revision 17.5U1-1998 #text_change 20.5un-2000 C.Accesion: A70807
Nature 393, 537-544, 1998 A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome	R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
A; Accession: F70806 A; Status: preliminary; nucleic acid sequence not shown; translation not shown A; Molecule type: DNA	Atauthors, 337, 344, 1338 A.Authorses, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A.Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A.Reference number: A70500; MUID:98295987; PMID:9634230
A;Residues: 1-1901 <col/> A;Cross-references: GB:AL022022; GB:AL123456; NID:g3261554; PIDN:CAA17745.1; PID:g292444 A;Experimental source: strain H37Rv	A;Accession: B70807 A.Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA
C;Genetics; A;Gene: Rv3508 C;Superfamily: collagen alpha 1(IV) chain	A; Residues: 1-1079 <col/> A; Cross-references: GB:AL022022; GB:AL123456; NID:g3261554; PIDN:CAA17749.1; PID:g292 A; Experimental source: strain H37Rv
Query Match 34.6%; Score 935; DB 2; Length 1901; Best Local Similarity 37.0%; Pred. No. 1.1e-43; Matches 264; Conservative 24; Mismatches 220; Indels 206; Gans 28;	C;Genetics: A;Gene: Rv3512 C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
TGGQQMGRIRIRGYGGLGGQGAGGAAAAAAAGGAG-QGGYGGLGSQGAGRGG 58	Query Match 34.3%; Score 927; DB 2; Length 1079; Best Local Similarity 44.3%; Pred. No. 1.9e-43; Matches 235; Conservative 15; Mismatches 202; Indels 78; Gaps 22;
59 QGAGAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAGGVGQGGLGGQ	Qy 7 GOQMGRIRIRGYGGLGGGGGGGGGGAGAAAAAAGGGCGGGGGGGGGGGGG
112 GAGGGAGAAAAAGGGGGGGGGGGGGGGGGGGGGGGG	Qy 67 AAGGAGOGGYGCLGSOGAGRGGLGGOGAGAAAAAGGVGQ 105
160	OY 106 GGLGGQGAGQBAAAAAAGGAGQGGYGGLGSQCAGRGGSGGQGAGAAAAAGGAG 161
198 GYGGLGGQGAQQGGYGGLGSQGAGRGGLGGQ-GAGAAA	Oy 162GGGGGGGSOGAGRGGLGGGGAGAAAAAAGGAGGGG 198
236 AGGAGGGLGGGGA GAAAAAAGGAGGGGGGGGGG	Oy 199 YGGLGGGGAGQGGYGGLGSQGAGRGLGGGGAGAAAAAGGAGGGGGGGGGAGA 251
274 RGGGGAGAAAAAGGA-GQGGYGGQGAGQGYGGLGSQGAG-RGGLGGQ-GAGAAA	OY 252 AAAAAGGAGGCGGGCGGGGGGGGGGAAAAGGAGGGGGGGG
327AAAAGGGAGGAGGAGGAGGAGGAGGAGGAGGAG	OY 310 QGAGRGCLGGOGAGAAAAAAGGAGOGCLGGOGAGAAAAAAGGGGCGCL 359
	Qy 360 GGQGAGQGAGAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQ-GAGAAAAAAGGAGGGGGG 418

```
hypothetical glycine-rich protein Rv3507 - Mycobacterium tuberculosis (strain H37RV) C;Species: Mycobacterium tuberculosis C;Species: Davin. 2000 C;Accession: E70806 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon S;Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Mature 393, 537-544, 1998 A;Ttle: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A;Reference number: A70500; MUID:98295987; PMID:9634230 A;Accession: E70806 A;Accession: E70806 A;Status: preliminary; nucleic acid sequence.not shown; translation not shown A;Molecule type: DNA A;Residues: 1-1381 ACOL>
A;Residu
     713 GAAGTGGTGGMIGTTGNAGVGGAGGSBGAGGTNGSGGAGGTDGQGGAGGAGGAGADNPTG 772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---AGRGGLGGQG--A 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGPGGTGGTGGNGGNGGNGGNGGNGGNGGNGGNAGNNSTNAPVGGEGGAGGDGGAGGAGG 620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55 ----GRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAA--AAAGGV-GQG 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLGGQGAGQGAGAAAAAAGG------AGQGGYGGLGSQGA------GRGGSGG 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGLGSQG-----AGRGGLGGQG------AGAAAAGG-AGQGGLGGQ-GAGAAAA 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GG----AGQGGYGGQG-AGQGGYGGLGSQG-----AGRGGLGGQ----GAGA 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----AGAA 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 ASWIGGOQMGRIRIRGYGGLGGQGAGQGAGAAAAAAGGAGQGGY-----GGLGSQGA 54
                                                                                                   --- AAGTGGTGGMIGTTGNA--
                                                   444 --GQGGYGGVGSGASAASAASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GAAAAAAGGA------GQGGYGGLGGQGA-------GQGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 33.0%; Score 892; DB 2; Length 1381; Best Local Similarity 34.7%; Pred. No. 1.8e-41; Matches 258; Conservative 25; Mismatches 218; Indels 242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QGAGAAAAAAGGA-----GQGGYGGLGSQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Superfamily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAAAAAAGGGGLGGQG-----
                                                                                                   773 IGGTGGDGGTGGAAGAGAGGGG--
                                                                                                                                                  502 QIGASNPGLSGCD 514
                                                                                                                                                                                                  --GVGGAGGQGGD 821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Rv3507
                                                                                                                                                                                                  811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             621
                                                     ò
                                                                                                 g
                                                                                                                                                  ð
                                                                                                                                                                                                음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                         C;Accession: 1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hanris, D.; Gordon, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A;Accession: D70807
A;Status: preliminary: nucleic acid semicon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Residues: 1-1489 <COL>
| Cross-references: GB:AL022022; GB:AL123456; NID:g3261554; PIDN:CAAL7751.1; PID:g292445
| Experimental source: strain H37Rv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34;
329 AAAGGAGGGGGG-GAGAAAAAGGAGQGGLGGQG-------AGQGAGA 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----GAGAAAAAA--GG----AGQGGYGGLGSQG-----AGRGGLGGQ-GAGAAAAAGGV 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GAGAAA----AAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAG 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -QGGYGGLGGQG-----AGQGGYGGLGSQGA----GRGGLGGQG----AG 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAAAGGGGGLGGQ-GAGAAAAAGGAGGGYGGLGSQGA------GRG 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGGAGGAAGTGGTGGMIGTTGNAGVGGAGGQGGDGGAGGAGADADQPGATGGTGFAGGAG 533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GO-GAGAAAAAAGGAGQGGYGGQGAGQCGYGGLGSQGA----GRGGLGGQG--AGAAAAA 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAAAAAAAGGV------RQGGYGGLGSQGA------GRGGQGAGAAAAAG 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GKAGGSSSAGGTNSSGSAGGTGRQSGTGGAGADNPTGIGGTGGDGGTGGAAGAGGAG 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -GAGGLGGQ-GAGAAAVGA---- 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 189;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1489;
                                                                                                      GLGGQGVGAGGLGGQGAGAAAAVGAGQGGYGGV-GSGASAASAASRLSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 33.8%; Score 912.5; DB 2; Best Local Similarity 39.1%; Pred. No. 1.5e-42; Matches 263; Conservative 19; Mismatches 202;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---GQGGLGGQGAGA-AAAAAGGAGQGGYGGLGSQG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A:Experimental source.
C:Genetics:
A:Gene: Rv3514
C:Superfamily: collagen alpha 1(IV) chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              G-AGGGGTGGL-
                       624
                                                                       419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 411
                       셤
                                                                       à
                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
```

32;

Gaps

348 AAAAGGAGGGLGGGGAGAAAAAAAAAAGGVRQGGYGGLGSQGA	Oy 363 GAGOGAGAAAAAAAAGGVROGGYGGLGSOGAGGAGAAAAAAGGAGOGGTGGL 420
DD 759 AFIGURAGGIGGSGGHAGAGGIGGAGGITITYFGNGGNAGGAGGNGGNGGNGGSGD 858 QY 492 LSSTISNVVSQIGASNPGLSG 512 i: : : Db 859 FGGNTTSGASGSGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNG	RESULT 9 A70869 hypothetical glycine-rich protein Rv2490c - Mycobacterium tuberculosis (strain H37RV)
	C.Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000 C.Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000 C.Accession: A70869 R.Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon; Connor, R.; Davise, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamilin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998 A.; Authors: Squares, S.; Barrell, B.G. A.; Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A.; Altitle: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A; Reference number: A70500; MUID:98395987; PMID:9634230
S. .G. olete genome	A.Accession: A70869 A.Status: Speliminary: nucleic acid sequence not shown; translation not shown A.Motacule type: DNA A.Residues: 1-1660 <col/> A.Residues: 1-1660 <col 1-1660="" <c<="" <col="" a.residues:="" td=""/>
	Query Match 32.2%; Score 868.5; DB 2; Length 1660; Best Local Similarity 41.5%; Pred. No. 3.9e-40; Matches 245; Conservative 22; Mismatches 175; Indels 149; Gaps 30;
Query Match 32.6%; Score 880; DB 2; Length 767; Query Match 32.6%; Pred. No. 5e-41; Matches 245; Conservative 22; Mismatches 218; Indels 96; Gaps 26;	OY 17 GYGGLGGOGAGOGGAGAAAAAGGGGGGGGGGGGGGGGGGG
GOGAGAAAAAAGGAGGGYGGLG-SQGAGRGCQ 59	OY 73 QGGYGGLGSQGA
60 GAGAAAAAAGGAGGGGGGGGGGGGGGG-GAAAAAGGVGGGLGGGGAGGGG 	Oy 103 -VGQGGLGGQG-AGQGAGA-AAAAAGGAGQGYGGLGSQGAGRGGSG 146 :
118 GAAAAAAGG-AGGGYGGLGSOGAGRGGSGG-OGAGAAAAAAGGAGGGGYGG 	OY 147 GOGAGAAAAAAGGAGGGGGGGGGGGGGGGGGGGAAAAAAGGGG
168 LGSQGAGRGGLGGQGAGAAAAAAAGGAGGGGYGGLGGQGAGQGGYG	QY 207 GOGG GYGGLGSOGAGRGGLGGGAGAAAAGGAGGGG 243 I
214 GLGSQGAGRGGLGGQGAGAAAAGGGGGGGGGGGGGAGAAAAAAGG LGGGGGGGGGG	OY 244 LGGOGA-GAAAAAAGGA-GOGGYGGLGSOGA-GRGGOGAGAAAAAGGA 289
AGQGGYGGLGSOGAGRGGQGAGAAAAAGA-GQGGYGGQGAGQGGYGG 	OY 290 -GQ-GGYGGGGAGQGGYGGLGSQGAGRGCLGGQ-GAGAAAAAAAAGGAGGGG 338
	339 LGGQGAGAAAAAGGAGQGGLGGQGAGQGAGAAAAAAAAGGVRQGGYGGLG
	Qy 392 SQGAGRGGQGAGAAAAAGGAGGGTGGLGGQGV 425

26;

63

150 468 191 524 643 347

```
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamiin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nathore 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A;Accession: A70896
A;Reference number: A70896
A;Reference number: A70896
A;Retaus: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA
A;Residues: 1-853 <COL>
A;Coss-references: GB:AL021897; GB:AL123456; NID:93256022; PIDN:CAA17207.1; PID:e125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hypothetical glycine-rich protein Rv0834c - Mycobacterium tuberculosis (strain H37RV) (Species: Mycobacterium tuberculosis (Species: Mycobacterium tuberculosis) (Mycobacterium tuberculo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A - - GGDGGDG - GNGCAGGTGGLLFCAGGNGGPGGSGGAADIGGNGGAGNGG - GTDGNGGN 699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           348 AAAAGGAGGGLGGQGAGQGA-----GAAAAAAAAAAAGGVRQGGYG--GLGSQG-- 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7 GOOMGRIRIRGYGGLGGOGAGOGAGAAAAAAAGGAGGGYGGLGSOGA---GRGGOGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGSGGHALLWGAGGAGGNGGSGGTGGAGGSTAGAGGNGGAGGGGGGGTGGLLFCNGGAGGHG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA-----GACAAAGGGGYGGLGSQGAGRGGLGGQ----GACAAAAAGGVGQGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGQGAGQGAGAAAAAAGGA----GQGGYGGLGSQG-----AGRGGSGGQGA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     525 GGAGGQAGGGGGNGCNG-GNGGNGGNGGNGATAGWLYGNGGAGGQGATAGAGGAGANGVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGQGGYGGQGAGQGGYGGLGSQGAGRGGLGGQ-GAGAAAAAAAGGAGQGGLGGQGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GESGGGAGSGGDGGGAGGAMLFGNGGAGGGGKGGNGAGGGLGGGSFGLPGLNGSGGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --GAAAAAAGGAGGGYGGLGSQGAGRGGLGGQGAGAA------AAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    469 GDGGAAAGTGIAGAGGRGGAG----GSGGSGGDGGGGAAGPAGWLFGDGGAGGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGAG--QGGYGGLGGQGAGQGGYGGLGSQGA-----GRGGLGGQGAGAAAAAGGA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----GQGGLGGQ-GAGAAAAAGGAGQGGYGGLG-SQGAG-RGGQGAGAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGRGGQGA--GAAAAAAGGAGQGGTGGLGGQGV--GAGGLGGQGAGAAAAVGAGQGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18; Mismatches 184; Indels 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 858; DB 2;
Pred. No. 8.4e-40;
                                                                                                                                                                                                                                                                                                                                                                                                                                             Gene: Rv1091; Superfamily: unassigned collagens
                                                                                                                                                                                                                                                                                                                                                                                A, Experimental source: strain H37Rv C, Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 232; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VG 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C;Comment: Much of the sequence consists of tandemly repeated 22-residue segments with C;Superfamily: Phascolus glycine-rich cell wall protein 1.8 C;Keywords: cell wall; structural protein; tandem repeat F;1-30/Domain: signal sequence #status pradicted <SIG> F;31-465/Product: glycine-rich cell wall protein 1.8 #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: DNA
A; Residues: 1-465 <KELD.
A; Cross-references: EMBL:X13596; NID:921002; PIDN:CAA31932.1; PID:921003
C; Comment: This protein is enriched in the cell wall fraction of young hypocotyls and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hypothetical glycine-rich protein Rv1091 - Mycobacterium tuberculosis (strain H37RV) C;Species: Mycobacterium tuberculosis (C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 24-Nov-1999 C;Accession: A70896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of
                                                                                                                                                                                                                                                                                                                                                                                                                                    Rikeller, B.; Sauer, N.; Lamb, C.J.
Embo J. 7, 3625-3633, 1988
Afille: Glycine-rich cell wall proteins in bean: gene structure and association A; Reference number: S01820; MUID:89091109; PMID:3208742
A; Accession: S01820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19;
                                   1184
                                                                                                                                                                                                                                                                                                            glycine-rich cell wall protein 1.8 precursor - kidney bean C; Species: Phaseolus vulgaris (kidney bean) C; Date: 30-Sep-1989 #sequence_revision 19-May-1994 #text_change 16-Jul-1999 C; Accession: S01820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GAGRGGLGGQGAGAAAAAGGVGQGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQCAGR 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -- AGGGYGTGGEHGGGYG--GGQGGG-GGYGGGGHGAAGYGGGEGGGGGGGGGGGA 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       82
   33 GYGLGHGTGGGYGGAAGSYGGGGGGGGGGGAAGEHGVYGYGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93 GAGYGGGGGSGGGGVAYGGGGE--RGGYGGGGG-----GGAG-GGYGAGGEHGIGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       194 AGGGGYGGLGGQGAGQ-GGYGGLGSQGAGRGGLGGQGAGAAAAAGGAGGAGGGGGGGGGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       317 GGG-GGGGGGAGGGYAAVGEHGGGYGGGGGG----DGGGYGTGGEHGGGYGGGQGGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23 GQGAGQGAGAAAAAAGGAGQGGYGGIGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGSGGQGAGAAAA-AAGGAGQGGYGGLGSQGAGRGG-----LGGQCAGAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  253 AAAAGGAGGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGQGGYGGLGSQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             313 GRGGLGGQGAGAAAAAAAGGAGGGLGGQGAGAAAAAAGGAGQGGL--GGQGAGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             142 GGGGGSGAGGGGGYNAGGAQGGGYGTGGGAGGGGGGGGDHGGGYGGGQGAGGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    371 AAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAG-GAGQGGTGGL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         53;
                                                                                                                                    Score 866.5; DB 1; Length 465; Pred. No. 1.8e-40;
                                                                                             ---GAGGLGGQGAGAAAVG-AGQGGYGG------VGSGASAASAAA 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 32.1%; Score 866.5; DB 1; Length 4 Best Local Similarity 47.8%; Pred. No. 1.8e-40; Matches 218; Conservative 8; Mismatches 177; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --GGQGVGAGGLGGQGAGAAAVGAGQGGYGGVGSG 454
                                                                                             426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                g
                                                                                          ò
                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      윱
```

```
31.5%;
36.5%;
                                                                                                                                                                                                                               Query Match 31.59
Best Local Similarity 36.59
Matches 241; Conservative
                                                                                                                                                                                  A,Gene: Rv0833
C,Superfamily: elastin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A 680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         491 A 491
                                                                                                                                                                                                                                                                                                                                                    58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    391
                                                                                                                                                                                                                                                                                                                      20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                     음
                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A; Reference number: A70500; MUID:98295987; PMID:9634230
A; Accession: B70812
A; Accession: B70812
A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Molecule: preliminary; nucleic acid sequence not shown; translation not shown
A; Residues: 1-882 <CCL>
A; Experiments: 1-882 <CCL>
A; Experimental source: strain H37Rv
C; Genetics:
A; Experimental source: strain H37Rv
C; Genetics:
A; Genetics:
C; Superfamily: elastin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               o, c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical glycine-rich protein Rv0833 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Species: 17-Ju1-1998 #sequence_revision 17-Ju1-1998 #text_change 20-Jun-2000
C;Accession: A7/0812
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
                                                                                                                                                                                                                29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        692 TGGAGGFGGAGAGIANFLGFSVLHGGAG-GAGGTATGTGGNGGAGGGGGLSSPVILGIGI 750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  487
                                                                                                                                                                                                                                                           213 NALMFGIGGNGGAGGAASGVGNGGVGGAGGAGGALVAIGGAGGAGGAATTGTGGAGGAGS 272
                                                                                                                                                                                                                                                                                                                                                                -----GQGGLGGQGA----GQGAGAAAAAGGA-------GQGGYGGLGSQGAG 141
                                                                                                                                                                                                                                                                                                                                                                                 RGGSGGQG------OGGY 165
                                                                                                                                                                                                                                                                                                                                                                                                                                            224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              383
                                                                                                                                                                                                                                                                                                       -----GAAAAAAAGGAGGGYGGLGSQG----AGRGGLGG---QGAGAAAAGG 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---AGRGGQGAGAAAAAAGAGQGGYGGQGAGQGGYGGLGSQGAG----RGGLGGQGAGA 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        632 ASSPFGIAIAIGGAGAQGGAGTSGATGGAGGDGVFEGIAVLGLGFGGAAGAGGAATGDGA 691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QGGYGGLGSQGAG------RGGQGAGAAAAAGGAGQGGTGGLGG-----QGVGA 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGAGGDGGGALGVLG-GMGGDGGDGGEAVAVGIAV-------GGAGGAGGAAPT 796
                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                            TGGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAGGA---GQGGYGGLGSQGAGRGGQGA 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LGGQGA-GAAAAAGGAGQGGLGGQGAGAAAAAAGGAG------QGGYGGLGSQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---GQGAGAAAAAAAAGGVR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAG-RGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGVGGAGGHGGGASVGTGGSGGAGGDGFGFVGAGGNGGNAGTGVGVNGANGGNGGSATGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA---AAAAGG---AGQGGLGGQ------GAGAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             572 LAAVGGAGAAGGDAISGTGGFGGAGGSARGLIFALGGAGAAGGDASTGVGGPGGPGGTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGLGGQGAGAAAAVGAGQGGYGGVGSGASAASAASRLSSPQASSRVSSAVSNLVASGPT
                                                                                                                                                                                                              209;
                                                                                                                                                                                    Length 882;
                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NSAALSSTISNVVSQIG-----ASNPGLSGCDVLI 517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GNGGAGGNGGDALGLVGVGGNGGNAGTGFGANTGGNGGDTII 838
                                                                                                                                                                                 31.8%; Score 857.5; DB 2;
35.2%; Pred. No. 9.2e-40;
Live 20; Mismatches 226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----AAAAGGAG-QGGLGGQGA------
                                                                                                                                                                                                              Conservative
                                                                                                                                                                                               Similarity
                                                                                                                                                                                                              247:
                                                                                                                                                                                  Query Match
                                                                                                                                                                                                Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    512
                                                                                                                                                                                                                                                                        157
                                                                                                                                                                                                                                                                                                                                                               103
                                                                                                                                                                                                                                                                                                                                                                                                                           142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        797
                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
```

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
Ajuthors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Reference number: A70812
A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Residues: 1-749 <COL>
A;Residues: 1-749 <COL>
A;Residues: 1-749 <COL>
A;Experimental source: GB:AL022004; GB:AL123456; NID:93261550; PIDN:CAA17639.1; PID:9291
A;Experimental source: strain H37Rv 25; 440 AGGILFGAGGAGGAVTGTGATGGSGGAGGGALLFGAGGAGGSSGIGGFAAGGAG 499 559 GSQG----AGRGGQGAGAAAAAGGAG------QGGTGGLGGQGVGAGGLGGQGA 435 193 267 QOALAAVNAPIQAVTGRPLIGNGANGAPGSGAPGGHGGWLFGGGGTGGSGVSGGAGGDGG 439 GAAAAVGA-GQGGYGGVG---SGASAASAASRL-SSPQASSRVSSAVSNLVASGPTNSA 490 620 DAGLLVGSGGVGGAGGFGLTTGGPGAAGDAGLFGSGGAGGAGGSGRTDLGGAGGAGGAG 679 200 LLALGDGGAGGAGTTGTGGAGGAGGKAGLLFGSGGAGGSGGAAGTFGDTGNSGGAGG 259 289 A------GQGGYGGQGGGYGGLGSQGA-----GRGGLGGQGAGAAAAAAAAGGA- 334 Gaps 79 7 GOOMGRIRIRGYGGLGGQ-----GAGQGAGAAAAAAGGA-GQGGYGGLGSQGAGRG 57 A------GOGGYGGLGGGGA-GOGGYGGLGSQGAGRG--GLGGGG--AGAAAAAGGAG AAAAGGAGQGGYGGLGSQG-AGRGGLGGQ------GAGAAAAA-----AGG 24.1 QGGLGGQGAGAA------AAAAGGAGQGGYGGL-----------GQGAGAAAAAAGGAGQGGYGGLGSQGAG------RGGLGGQGAGAAA----------GRGGGGGGAAAAAGG 21; Mismatches 223; Indels 176; Length 749; Score 850; DB 2; Pred. No. 2.1e-39;

E70917
hypothetical glycine-rich protein Rv1450c - Mycobacterium tuberculosis (strain H37RV)
C; Species: Mycobacterium tuberculosis
C; Species: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

œ

```
327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          연
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
C;Accession: E70917
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, E.; Connor, R.; Davies, R.; Devilin, K.; Feltwell, T.; Gentles, S.; Hamilin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Hamilin, N.; Holroyd, S. Nature 393, 537-544, 1998
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Authors: Sqares, R.; Sulston, J.E.; Paylor, K.; Whitehead, S.; Barrell, B.G.
A;Accession: E70917
A;Accession: E70917
A;Accession: E70917
A;Kesidues: prealiminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1329 <COL>
A;Accession: E70917
A;Residues: 1-1329 <COL>
A;Accession: E70917
A;Residues: 1-1319 <COL>
A;Accession: E70917
A;Residues: 1-1319 <COL>
A;Accession: E70917
A;A
                                                                                                                                                                                                                                                                                                                                                                                                 C;Genetics:
A;Gene: Rv1450c
C;Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical glycine-rich protein Rv2634c - Mycobacterium tuberculosis (strain H37RV) C; Species: Mycobacterium tuberculosis (C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 24-Nov-1999 C; Accession: F70963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                644 GAGGKGGDGMAGVAGSSPTTAGESGTSGQNGGAGGAGGAGGRGGDFGGDGTGGAGGNGA 703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SQG--AGRGGQ-----GAGAAAAAGGAGQGGTGGLGGQGVG----AGGLGGQGAGAAAA 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        409 SGGKAGGNGGAGGAGGLVGNGGAGGAGGNGAPGAPPSGGDPNGGGGGAGGAG-GKGGDG- 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GAAAAAAGGAGGGGYGGLGSQGA-GRGGLGGQGAGAAAAAAGGVGQGGLGGQGAGQGAG-- 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------AAAAAAGGAGQGCYGGLGSQGA-GRGGSGGQGAGAAAAAGGAGQG 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     525 GGKALAATHQDGSMGAGGAGGNGGAGGMGGDGGNGAKGTFDNGGDGVGGNGGNGGSRGIG 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAG-----GAGQGGYGGQG-----AGQGGYGGLGS----QGA-----GRG 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        823 GDGGNGGNSQVVGGNGGDGGNGGNGGSAGTGGNGGRGGDGAFGGMSANATNPGENGPNGN 882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PGGNGGAGGAGGAGLNGGNGGAGGNGGLGGFGGNGAAGANGVAVGAPGQPGGAGGHGGAG 942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----GAAAA 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TGGQQMGRIRIRGYGGL---GGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGA 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GQGGYGGLGSQG-----AGRGGLGGQ--GAGAAAAGG-----AGQGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GYGGLGSQGA-------GRGGLGGQGAGAAAAAAAGGA-GQGGYGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 31.4%; Score 847; DB 2; Length 1329; Best Local Similarity 38.9%; Pred. No. 4.8e-39; Matches 243; Conservative 16; Mismatches 196; Indels 170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          250 GAAAAAGGA--GQGGYGGLGSQGAGRGGQGA-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLGGQG-------AGAAAAAAAAAGGAGQGGLGGQGA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GNAGAGGPGGTGSTAGKAGPAGSIL 1027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                441 VGAGQGGYGGVGSGASAASAASRL 465
                                                                                                                                                                                                                                                                                                                                                                            A; Experimental source: strain H37Rv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno A;Reference number: A70500; MUID: 98295987; PMID: 9634230
A;Coession: F70963
A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA
A;Coss-references: GB: Z80225; GB: ALL23456; NID: 93242265; PIDN: CAB02341.1; PID: e26639
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv2634c
C;Superfamily: unassigned collagens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GAAAAAAAGGAGGGGYGGLG-----GQGAGQGGYGGLGSQG-----AGRGGLGGQGAGA 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YGGLGSQG--AGRGGQ--GAGAA-----AAAAGGAGQGGT---GGLGGQGVGAGGLGGQ 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 ASMTGGQ-----QMGRIRIRGYGGLGGQGAGQG--AGAAAAAAGGA-GQGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       207 ATLVGGTGGVGGATGLIGSGGFGGAGGAAGVGTTGGVGGSGGVGGVFGNGGFGGAGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  103 -VGQGGLGGQG------AGQGAGAAAAAGGA-GQGGYGGLGSQGAG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              565 GAACFAGGV - GGAGGEGLIDGAGTAEGGTGGLGGVGGTGGMGGSGGVGGNGGAAGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 387 NPNTGANGTDNSGNGNQTGGNGGPGPAGGYGEAGGY--GGQGGLGESLDGNDGTGGKGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             445 GGTAGTDGGAGGAGGAGGIGETDGSAGGVATGGEGGDGATGGVDGGVGGAGGKGGQGHNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAAAGGAGQGGLGCQGAGAAAAAAGGAGQGGYGGLGSQG------AGRGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAAAAAGGAGGGYGGQG-----AGQGGYGGLGSQG----AGRGGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 AGRGGGGAGAAAAAGGAGGGYGGLGSQGA-----GRGGLGG-QGAGAAAAAGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :| || || || || || || 327 LLGDGGAGGGGGGGGAVAGVLGGMPGAGGNGGNANWFGSGGAGGGGGTGLAGTNGVNPGSIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LIGLGGGGGAGGVGGTG-----GIGGIGGAGGNGGAGGAGTTTGGGATIGGGGTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24; Mismatches 204; Indels 126;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 778;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GA----GAAAAVGAGQGGYGGVGSGASAASAAASRLSSP 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ouery Match 31.3%; Score 844; DB 2; Best Local Similarity 38.9%; Pred. No. 4.5e-39; Matches 225; Conservative 24; Mismatches 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GATGGGGGDFALG-GNGGAGGAGGSPGGSSGIQGNMGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search completed: December 18, 2002, 16:22:22 Job time : 21.6009 secs
```

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                          OM protein - protein search, using sw model
```

December 18, 2002, 16:06:04; Search time 8.10437 Seconds (without alignments) 2702.183 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-490-291-8 2700 1 MASWIGGQQMGRIRIRGYGG......GLSGCDVLIQALLGHHHHH 528

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

112892 segs, 41476328 residues Searched:

112892 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_40:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		æ			SUMMAKIES		
Result		Query					
. No	Score	Match	Length	BB :	ΩI	Description	
7	2273	84.2	747	Н	SPD1_NEPCL		la cla
7	1148		627	-	SPD2_NEPCL		
e	952.5	35.3	5263	Н	FBOH_BOMMO		bombyx mori
₹	935	4	1901	Н	YZ08_MYCTU	_	acteri
S	866.5	S	465	Н	GRP2_PHAVU		phaseolus v
9	844	31.3	778	Н	YQ34_MYCTU		acteri
7	839.5	31.1	914	Н	WA22_MYCTU		acteri
80	828.5	30.7	641	Н	EBN1_EBV		epstein-bar
σ	819.5	30.4		-	Y747_MYCTU		acteri
10	807	29.9	957	Н	Y278_MYCTU	_	mycobacteri
11	782.5	29.0	603	Н	YD25_MYCTU	_	mycobacter1
12	736	27.3	384	Н	GRP1_PETHY		petunia hyb
13	729.5	27.0	860	Н	ELS_MOUSE		musculu
14	714.5	26.5	338	Н	GRP_ARATH		arabidopsis
15	697.5	25.8	481	Н	LORI_MOUSE		musculu
16	684	25.3	543	Н	YP91_MYCTU	_	mycobacteri
17	680.5	25.2	864	Н	ELS_RAT	099372 rattus	s norv
18	671.5	24.9	498		YI18_MYCTU	Q50615 mycobacter	acteri
19	612	22.7	463	Н	YA68_MYCTU	_	mycobacteri
20	609.5	22.6	515		YI40_MYCTU		acteri
21	602	22.3	491	Н	YK98_MYCTU		mycobacteri
22	595	22.0	747	-	ELS_BOVIN		aurus
23	593.5	22.0	750	-	ELS_CHICK	PO7916 gallus	gallus gall
24	536.5	19.9	730	Н	ELS_HUMAN	homo	sapien
25	526.5	19.5	252	Н	GRP1_PHAVU	phase	phaseolus v
56	516	19.1	1733	-	VNUA_PRVKA	_	pseudorabie
27	512.5	19.0	672	-	PHX5_MOUSE		musculu
28	501.5	18.6	1356	Н	CA21_ONCMY		oncorhynchu
29	471.5	17.5	1355	-	CA21_RANCA		catesb
30	470	17.4	2038	Н	FSH_DROME	P13709 drosophila	phila
31	455.5	16.9	996	-	FIB1_PETMA	petro	myzon
32	454	16.8	316	~	LORI_HUMAN	P23490 homo s	sapien
33	439	16.3	671	Н	CA11_RAT	rattn	s norv

P30754 riftie pach O76743 caenorhabdi P02457 gallus gall P02451 homo sapien P02452 homo sapien P11087 mus musculu P28481 mus musculu O9xsj7 canis famil P02458 homo sapien P02458 homo sapien P08123 homo sapien
CAFF_RIFPA GLH4_CAEEL CA11_CHICK CA11_CHICK CA11_HUMAN CA11_MOUSE CA11_CANFA CA21_CANFA CA21_CANFA CA21_CANFA CA21_CANFA
аннаннанна
1027 1156 11465 14664 14664 11460 11372 11316 11418
100.00 100.00 100.00 100.00 100.00 100.00 100.00
434 432 429.5 428.5 427 427 427 419 419 411.5
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

## ALIGNMENTS

~

```
15;
                                                                                                                                                                                                                                                                                                                                          S -> T (IN REF. 1).
NPGLSGCDVLIQALLEVVSALIQILGSSSIGQVNYGSAGQA
TQ1VGQSVYQAL -> ILVFLDVMSSFKLFSRLFLLLSRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAAGGAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQGGLGG---QGAGQGA-GA 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---GAGGGGYG 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGQQMGRIRIRGYGGLGGQGAGQAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 498 AAAVGAGQEGIRGQGAGQGGYGGLGSQGSGRGGLGGQGAGAAAAAAGAGGGGGGGGGGGGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -------AAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAAAGGAGGGLG----GQGAGAAAAAAGGAGGGYGGLGSQGAGR---GGQGAGA-AAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAGGAGGGYGGLGSQGAGRGCQGAGAAAAAGGAGQRGYGGLGNQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA------AAAAAGGAGO------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---GELGGQGAGAAAAAGGAGQGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QGAGAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLGGQGAGQGAGASAAAAGGAGQGGYGGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10; Indels 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 747;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLGSQGAGRGGLGGQGAGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                  850E44B0D649E012 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          aaaaaggacoccoccoccoccoccoccoccoccoccoccoccc
                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1;
                                                                                                                                                                                                                                                                                                                                L (IN REF. 1)
T (IN REF. 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                9.1e-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 9.1e
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 2273;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4;
                                                                                                                                                                                                                                                                                                                                                                                                                  60528 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                84.2%;
illarity 71.0%;
Conservative
                                                                                                    2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

2003

                                                                                                                                                                                                                                                                                                                                                                                                                  747 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
494;
                                                                                                                                                                                                                                                                                            REPEAT
REPEAT
CONFLICT
                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                   REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                      REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                            REPEAT
REPEAT
REPEAT
                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              614
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ç
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 30;
                                                                                                                                                                                                                                                                                            ----GGQGAGAAAAAAGGAGGGYGGLGSQGAGRGGLG-GQ----GAGAAAAAGGV---- 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAAAAAGPGGYGPGQQGPGGYGPGQQGPGRYG 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GYG----GLGGQGAGQ----GAG-AAAAAAGGAGQGGYG-----GLGSQGAGR--
                                                                                                                                                                 Eukaryota; Metazoa; Arthropoda; Chellcerata; Arachnida; Araneac;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 136;
                                                                                                                                                                                                                                      Hinman M.B., Lewis R.V.;
"Isolation of a clone encoding a second dragline silk fibroin.
Nephila clavipes dragline silk is a two-protein fiber.";
J. Biol. Chem. 267:19320-19324(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          627;
                                                                                                                                                                                                                                                                        J. Biol. Chem. 267:19320-19324(1992).
-I- FUNCTION: Spiders major ampullate silk possesses unique
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15 APPROXIMATE TANDEM REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CB9B63779B2C594B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       42.5%; Score 1148; DB 1;
51.1%; Pred. No. 1.7e-46;
1ve 25; Mismatches 139;
                                                                                                      01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Spidroin 2 (Dragline silk fibroin 2) (Fragment).
Nephila clavipes (Orb spider)
            521
                                                                                  ¥.
PTNSAALSSTISNVVSQIGASNPGLSGCDVLIQALL
                                                                                  627
                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE~92406876; PubMed=1527052;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8.
10.
11.
112.
14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M92913; AAA29381.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      516
627 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                         NCBI_TaxID=6915;
                                                                                                                                                                              Araneomorphae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Sim
Matches 314;
                                                                                SPD2_NEPCL
P46804;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
REPEAT
                                                       RESULT 2
SPD2_NEPCL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
486
                     674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57
                                                                                             ŏ
                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
```

us-09-490-291-8.rsp

```
GLGGGGGAAAAAAAGGGGGLGGGGAG-------AAAAAAG----GAGGGG 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            123 SGQGGPGGYGPGO------QGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAA 171
                                                                                                                                                                                                                                                                                                                                                                                     261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGYG----GLGSQGAG-RGGQGAG-AAAAAGGAGQGGYGGQGAGQGGYGGLGSQGAGRG 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   343 GPGGYGPGSASAAAAAAGPGQQGPGGYGPGQQGPSGPGSASAAAAAAAAAAAGPGGYGPGQQG 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Tsujimoto Y., Suzuki Y.; "The DNA sequence of Bombyx mori fibroin gene including the 5' flanking, mRNA coding, entire intervening and fibroin protein coding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pancrustacea; Hexapoda;
Lepidoptera; Glossata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -GQGGLGGQGAGAGAAAAAGGAGQGGYG----GLGSQGAGRGGSGGQG-AGAAAAAA
                                                                                                                                                                                                                                                                                                                               158 GGAGQ---GGYG----GLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGGQGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GYGGLGSQGAGRGGLGGQGAGAAAAAGGAGQGGLGGQGAG------AAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    286 GGYGPGQQGPGGYGPGQQGPSGAGSAAAAAAAGPGQQGLGGYGPGQQGPGG---YGPGQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAGGAGQGGTG----GLGGQGV-----GAGGLG------GQGAGAAAAVGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    460 AAAAAGPGGYGPAQQGPSGPGIAASAASAGPGGYGPAQQGPAGYGPGSAVAASAGAGSAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             448 YGGVGSGASAASAASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE-20330362; PubMed-10871375;
Zhou C.-2., Confalonieri F., Medina N., Zivanovic Y., Esnault C.,
Yang T., Jacquet M., Janin J., Duguet M., Perasso R., Li Z.-G.,
Fine organization of Bombyx mori fibroin heavy chain gene.";
Nucleic Acids Res. 28:2413-2419(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FBOH_BOMMO STANDARD; PRT; 5263 AA. P05790; 026379; 017220; 017020; 0100v-1988 (Rel. 09, Created) 16-00T-2001 (Rel. 40, Last sequence update) 16-00T-2001 (Rel. 40, Last annotation update) 15-00T-2001 (Rel. 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Arthropoda; Mandibulata;
Insecta; Pterygota; Neoptera; Endopterygota;
Ditrysia; Bombycoidea; Bombycidae; Bombyx.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-168 FROM N.A. MEDLINE=80045039; PubMed=498286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PARTIAL SEQUENCE FROM N.A.
MEDLINE-79211211; PubMed-455439;
Tsujimoto Y., Suzuki Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bombyx mori (Silk moth).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              508 PGLSGCDVLIQALL 521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cell 18:591-600(1979).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=7091;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           regions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       408
                                                                                                                                         104
                                                                                                                                         à
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1D DTT DDT ACC DDT ACC
                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

```
Tanaka K., Kajiyama N., Ishikura K., Waga S., Kikuchi A., Ohtomo K., Takagi T., Mizuno S.;
Takagi T., Mizuno T., Mizuno T., Mizuno S.;
Takagi T., Mizuno T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way mon-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PARTIAL SEQUENCE FROM N.A. MEDINE-9435842; Pubmed-7916056; MEDINE-9435842; Pubmed-7916056; Mita K., Ichimura S., James T.C.; "Highly repetitive structure and its organization of the silk fibroin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 35.3%; Score 952.5; DB 1; Length 5263; Best Local Similarity 43.3%; Pred. No. 5.7e-37; Matches 249; Conservative 28; Mismatches 201; Indels 97; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17 GYGGLGGQGAGQGAGAAAAAAAG-GAGQGGYGGLGS-----QGAGRG-----GQGAGAAA 65
                                                                                                                                                                                                                the secondary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INTERCHAIN (WITH LIGHT CHAIN).
        end and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10 C -> V (IN REF. 2).
391586 MW; 8EE11D3AOA47440E CRC64;
                                                                                                                               STRAIN-Kinshu X Showa;
MEDI.INE-89094868; PubMed-3210244;
Mita K., Ichimura S., Zama M., James T.C.;
Mita Codon usage pattern and its implications on structure of slik fibroin mRNA.";
J. Nol. Biol. 203:917-925(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 5179-5263 FROM N.A., AND DISULFIDE BONDS
     the fibroin gene at the 5'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FIBROIN HEAVY CHAIN. HIGHLY REPETITIVE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-J-139;
MEDLINE-99296390; PubMed-10366732;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AF226688; AAF76983.1; --
EMBL; V00094; CAA23432.1; --
EMBL; V00097; CAA23433.1; --
EMBL; S74439; AAB31861.1; --
EMBL; X13869; CAA32076.1; --
EMBL; M35378; AAA27839.1; --
EMBL; M35378; AAA27839.1; --
EMBL; S01844; S01844.
S11K; S19nal; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mol. Evol. 38:583-592(1994).
                                                                                                      PARTIAL SEQUENCE FROM N.A.
"Structural analysis of
surrounding regions.";
Cell 16:425-436(1979).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5263
5206
5244
5263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5263 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5244
5260
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
```

```
SEQUENCE
                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 5
GRP2_PHAVU
                                                                                                                                                                                                                                                              29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     962
                                                                                                                                                                                                                                                                                                          112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              426
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 K 1
                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration detween the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                          2868
                            2927
                                                                                                                                                                     2987
                                                                                                                                                                                                                                                             3107
                                                                                                                                                                                                                                                                                                                                           323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                'n
                                                                                                                                              275
                                                                                                                                                                                                                                       -----GLG-GQGAG 365
                                                                                                                                                                                                                                                                                   421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cole S.T., Brosch R., Parkhill J., Garnler T., Churcher C., Harris Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Deviln K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Hamlin N., Holroyd S., Osborne J., Quall M.A., Rajanfram M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; Deciphering the biology of Mycobacterium tuberculosis from the Nature 393:537-544(1998).
                                                                  GQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAAGGAGQGGYGGLGSQGA
                                                                                                GRGGLGGQGAGAAAAAAG----GAGQG-GYGGLGGQGAGQGGYGGLGSQGAGRGGLGG
                                                                                                                                            QGAGAAAAAG----GYGGLGGQGAGAAAAAAGGAGQG------GYGGLGSQGAGRG
                                                                                                                                                                                         QGAGAAAAAAAAAGGVRQGGYGGLGSQGAGRG-GQGAGAAAAAAAGGAGQGGTGGLG---
                                                                                                                                                                                                                                                                                                                             --GQGVGAG-----GLG------GQGAGAAAAVGAGQGGYGGVGSGASAASAAS
        AAAGGAGGGYGGLGS-----QGAGRGGLGGQGAGAAAAG-----GVGQGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
116-0201 (Rel. 40, Last annotation update)
Hypothetical PE-1885 family protein Rv3508 precursor.
Rv3508 OR MTV023.15.
                                                                                                                                                                                                                                                                                                                                                                                                3227 GYGAGAGSGAASGAGAGAGAGAGTGSSGFGPYVAN 3261
                                                                                                                                                                                                                                                                                                                                                                         RLSSPQASSRVSSAVSNLVASGPTNSAALSSTISN 498
                                                                                                                                                                                                                                     AAAAAAAGGAGGGLGGQGAGAAAAAAGGAGQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 1901 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-H37Rv;
MEDLINE-98295987; PubMed-9634230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      YZ08_MYCTU
O53553;
                                                                                                174
                                                                                                                     2869 (
                                                                                                                                                                  2928
                                                                                                                                                                                                                                                                                                                                                                          464
                                                                                                                                                                                         276
                                                                                                                                                                                                                                                            3048
       99
                                                                                                                                            228
                                                                                                                                                                                                                                                                                 366
                                                                                                                                                                                                                                                                                                        3108
                                                                                                                                                                                                                                     324
                                                                                                                                                                                                                                                                                                                              422
                                                   114
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                        엽
                                                                       셤
                                                                                                ö
                                                                                                                    셤
                                                                                                                                         ò
                                                                                                                                                                g
                                                                                                                                                                                        ò
                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                         à
        õ
                                                   ç
```

```
modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                               28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | ||||
1081 nagygltakagdggaaggaggaggaggaggggggggggggggssttsin 1140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1201 VGGDGGEGASGLGLGLSGFDGGGGGGGGGAGGSAGAGGINGAGGAGGTGGAGGDGAPATLI 1260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1261 GGPDGGDGGQGGIGGDGGNAGFGAGVPGDGGNAGFGAGVPGDGGIGGTGGAGGAGGA 1320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         198 GYGGLGGQG----AGQGGYGGLGSQG------AGRGGLGGQ-GAGAAAA 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          781
                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 TGG0--QMGRIRIRGYGGLGGQGAGQGAGAAA---AAGGAG-QGGYGGLGSQGAGRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QGAGAAAAAAGGAGGGYGGLGSQGAGR-----GGLGGQGAGAAAAAGGVGQGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                842 AGGNGGDGGDGATGAAGLGDNGGVGGDGGAGGAAGNGGNAGVGLTAKAGDGGAAGNGGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          902 GAGGAGGAGDNNFNGGOGGAGGOGGLGGASTTSINANGGAGGNGGTGGKGGAGGAGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               236 AGGAGQGGLGGQGA-----GAAAAAGGAGQAGGYGGLGSQGA-----G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RGGQGAGAAAAAAGGA-GQGGYGGQGAGQGGYGGLGSQGAG-RGGLGGQ-GAGAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                723 GGTNGSGGAGGAGGKGGTGGAGGSGADNPTGAGFAGGAGGTG-GAAGAGGAGGATGTGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GAGQGAGAAAAAAGGAGQGGYGGLGSQGAG-----RGGSGGQGAGAAAAAAGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1141 ANGGAGGNGGTGGKGCAGGAGTLGVGGSGGTGGDGGDAGSGGGGGFGGAAGKAGGGGCNGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --------AGAAAAGGA------GQGGLGGQG----AGAAAAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YGGLGSQGA-----GRGGQGAGAAAAAAGG----AGQGGTGGLGGQGV----
                                                                                                                                                                                                                                                                                      HYPOTHETICAL PE-PGRS FAMILY PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GADGDPSIDGGQGGAGGHGGQGGKGGLNSTGLASAASGDGGNGGAGGAGGNGGD 1374
                                                                                                                                                                                                                                                                                                                                                                                       34.6%; Score 935; DB 1; Length 1901; 37.0%; Pred. No. 1.8e-36; Live 24; Mismatches 220; Indels 206;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        461 AASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCD
                                                                                                                                                                                                                                                                                                                 RV3508.
W; C7B1923D5D0146CD CRC64;
                                                                                                                                                                                                                                      Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      160 -AGQGGYGGLGSQGAG----RGGLGGQ-GAGAAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         465
                                                                                                                     Tuberculist; Rv3508; -...
InterPro: IPR000084; PE_region.
Pfam; PF00934; PE; 1.
Hypother PD00123; PE_region; 1.
Hypothetical protein; Repeat; Signal; C.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----GAGGLGGQGAGAAAVGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                       1901 AA; 147627 MW;
                                                                                                EMBL; AL022022; CAA17745.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GRP2_PHAVU STANDARD;
P10496;
01-JUL-1989 (Rel. 11, Created)
                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 37.08
Matches 264; Conservative
                                                                                                                                                                                                                                                                                 31 1901
```

Ŋ

```
ð
                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19;
                                                                                                                                                           01-JUL-1989 (Rel. 11, Last sequence update)
01-OCT-1994 (Rel. 30, Last annotation update)
61yGine-rich cell wall structural protein 1.8 precursor (GRP 1.8).
Phaseolus vulgaris (Kidney bean) (French bean).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnollophyta; eudlcotyledons; core eudicots; Rosidae;
eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GRGGLGGQGAGAAAAAAAGGAGQGGLGGQGAGAAAAAAAGGAGQGGL--GGQGAGQGAGA 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    317 GGG-GGOGGGAGGGYAAVGEHGGGYGGGOGGG----DGGGYGTGGEHGGGYGGGOGGG- 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GAGRGGLGGQGAGAAAAAGGVGQGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGR 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           143 GGSGGQGAGAAA-AAGGAGQGGYGGLGSQGAGRGG-----LGGQGAGAAAAAAGG 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      194 AGGGGYGGLGGQGAGQ-GGYGGLGSQGAGRGGLGGQGAGAAAAGGAGGAGGGLGGQGAGAA 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       253 AAAAGGAGGGGGGGGGSQGAGRGGQGAGAAAAAGGAGQGGYGGQGAGQGGYGGLGSQGA 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     371 -- AGGGYGTGGEHGGGYG--GGQGGG-GGYGAGGDHGAAGYGGGEGGGGGGGGGGGGAGDGGA 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23 GQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQ 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GYGLGHGTGGGYGGAAGSYGGGGGGGGGGGGAAGEHGVVGYGGGGSGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                            SHEET CONFIGURATION.
SIMILARITY: THE N-TERMINAL SIGNAL SEQUENCE OF ABOUT 30 AA SHOWS
ABOUT 60% HOMOLOGY TO THAT OF THE GLYCINE-RICH CELL WALL
STRUCTURAL PROTEIN GRP 1.0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     142 GGGGGSGAGGGGGYNAGGAQGGGYGTGGCAGGGGGGGGDHGGGYGGGQGAGGGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAG-GAGQGGTGGL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 1.8. GLY-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 866.5; DB 1; Length 465; Pred. No. 9.5e-34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8; Mismatches 177; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8 X 22 AA TANDEM REPEATS.
B5C4A9B983B43607 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; S01820; S01820.
Cell wall; Structural protein; Repeat; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         33 465 G
205 359 8
465 AA; 36683 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X13596; CAA31932.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         47.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 218; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                 NCBI_TaxID=3885;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      313
     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content. Is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
STRAIN=CDC 1551 / Oshkosh;
STRAIN=CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weldman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                          <u>.</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                 Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBL_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                            Harris
                                                                                                                                                                                                                                                                                                                                                                                                                    cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris Gordon S.V., Elgimeler K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Deviln K., Feltwell T., Gentles S., Hamin N., Holroyd P., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Seeger K., Krogh A., McLean J., Moule S., Murphy L., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J. E., Taylor K., Whitehead S., Barrell B.G.; "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels 126;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 778;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hypothetical protein; Complete proteome.

CONFLICT 51 51 V -> L (IN REF. 2).

CONFLICT 63 63 Q -> H (IN REF. 2).

CONFLICT 63 74 A -> T (IN REF. 2).

SEQUENCE 778 AA; 63131 MW; DAB20PE58E4999E7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Whole genome comparison of Mycobacterium tuberculosis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31.3%; Score 844; DB 1; 38.9%; Pred. No. 1.4e-32;
                   01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv2634C.
RV2634C OR MT2712 OR MTCY441.04C.
Mycobacterium tuberculosis.
454
                                                                                                                                  778 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24; Mismatches
 --GGQGVGAGGLGGQGAGAAAAVGAGQGGYGGVGSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AE007103; AAK47026.1; ALT_INIT
                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-98295987; PubMed-9634230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tuberculist; Rv2634c; -.
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE; 1.
ProDom; PD001223; PE_region; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  274 2/4
778 AA; 63131 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; 280225; CAB02341.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               complete genome sequence
Nature 393;537-544(1998)
                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                               [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TIGR; MT2712;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Sim
Matches 225;
                                                                                                                                YQ34_MYCTU
P71933;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bisha1 W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                             RESULT 6
YQ34_MYCTU
                                     426
421
                                   g
```

```
Bishai W.;
                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             392
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-9825987; pubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eaglameler K., Gas S., Barry C.E., III, Frela F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Woule S., Murphy L.,
Oliver S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
                                                                      267 AAGGVGGAASYFGTGGGGGVGGDGAPGGDGGAGPLLIGNGGVGGLGGAGAAGGNGGAGGM 326
                                                                                                                                                 233 AAAAGGAGGGGGGGAGAAAAAAGGAGQGGYGGLGSQG------AGRGGQGAG 280
                                                                                                                                                                                                                                                                            564
                                                                                                                                                                                                                                                                                                                   565 GAAGFAGGV--GGAGGEGIIDGAGTAEGGIGGLGGLGGLGGGGGGGGGGGGGGGGGGGAGGS 622
                                                                                                                                                                                                                                                                                                                                                                                                     433
                                                                                                                                                                         387 NPNTGANGTDNSGNGNOTGGNGGPGPAGGVGEAGGV--GGOGGLGESLDGNDGTGGKGGA 444
                                                                                                                                                                                                                                                                                                  281 AAAAAAGGAGGGYGGQG-----AGQGGYGGLGSQG----AGRGGLGGQGAGAAA 326
                                                                                                                                                                                                                                                                                                                                                 674
                                                                                                                                                                                                                                                                                                                                                                                                                          675 VGGAGGTGGTGGAGGTTGGSGGAGGLIGWAGAAGGTGAGGTGGQGGLGGQG-GNGGNGGT 733
  53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.NCBI_TaxID=1773;
                                              54 AGRGGQGAGAAAAAAGGAGQGGYGGLGSQGA-----GRGGLGG-QGAGAAAAAGG----
ASMTGGQ-----QMGRIRIRGYGGLGGQGAGQG--AGAAAAAAGGA-GQGGYGGLGSQG
                       207 ATLVGGTGGVGGATGLIGSGGFGCAGCAAAGVGTTGGVGGSGGVGGVFGNGGFGGAGGLG
                                                                                                                        327 LLGDGGAGGGGGPAVAGVLGGMPGAGGNGGNANWFGSGGAGGOGGTGLAGTNGVNPGSIA
                                                                                                                                                                                                 103 -VGQGGLGGQG------AGQGAGAAAAAAGGA-GQGGYGGLGSQGAG----
                                                                                                                                                                                                                                                                505 GVGDAFGGDGGIGGDGNGALGAAGGNGGTGGAGGNGCRGGMLIGNGGAGGAGGTGGTGGTGGG
                                                                                                                                                                                                                                                                                                                                                                         YGGLGSQG--AGRGGQ--GAGAA-----AAAAGGAGQGGT---GGLGGQGVGAGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN-CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eiscn J.A., Carpenter L., White O.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                  GA----GAAAAVGAGQGGYGGVGSGASAASAAASRLSSP 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    734 GATGGGGGDFALG-GNGGAGGSPGGSSGIQGNMGPP 771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
MAG22 antigen precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        914 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-H37RV;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WA22_MYCTU
006794;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
WA22_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                  387
                                                                                                                                                                                                                                                                                                                                                                                                                                                  434
                                                                                                                                                  142
                      g
                                                                      g
                                                                                                                       쉽
                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                ò
                                                                                                 ð
                                                                                                                                                à
                                                                                                                                                                                                 ð
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                  ç
                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29;
                                                          Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     129 GGDAGWLIGNGGAGGSGAKGANGGAGGPGGAAGLFGNGGAGGAGGTATANNGIGGAGGAG 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAAAGGV----GQGGLGGQGAGQAGAAAAA---GGAGQGGYGGLGSQGA-GRGGSGGQ 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GAGAAAAAAGGAGGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAG------QGG 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          486 FGGFGGLLLGADGYNAPESTSPWHNLQQDILSFINEPTEALTGRPLIGNGDSGTPGTGDD 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----SQGAGRGGQGAGAAAAAGGAG-----QGGTGGLGGQG--VGAGGLGGQGA- 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                    "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6 GGQQMGRIRIRGYGGLGGQGAGQAGAAAAA-----AGGAGQ--------GGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SQ----GAGRGGGGAGAAAAAGG----AGQGGYGGLGSQGAGRGGLGG-----QGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         309 GGGGGLFGAG--GDGGAGGSGLTTGGAAGNGG-NAGTLSLGAAGGAGGTGGAGGTVFGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          199 YGGLGGQGA-----GQGGYGGLGSQGAGRGGLGGQGAGAAAAAGGAGQGGLGGQG--AG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     366 KGGAGGAGGNAGMLFGSGGGGGTGGFGFAAGGQGGVGGSAGMLSGSGGSGGAGGSGGPAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAAAAAGGA------GQGGYGGLGSQGAGRGG-QGAGAAAAAAGGAGQGGYGGQG--AG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------GSQGAG-R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGLGG------QGAGAAAAAAAGGA--------GQGGLGGQGAGAAAAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          || || || GGAGGWLFGNGGNGGAGGAGGAGGTLFGTGGAGGAGGVGTAGAGGAGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GA----GQGGLGGQGAGQ------GAGAAAAAAAAAG-------GVRQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SAFLIGSGGTGGVGGATTTGGVGGAGGNAGLLIGAAGLGGCGGGAFTAGVTTGGAGGTG
                                  Dodson R., Gwinn M.L., Haft D., Hickey E.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels 185;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                914;
                                                                                                                                                                                                                                           PGRS
                                                                                                                                                                                                                                                                                            -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 85.
                                                                                                                                                                          laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       914 WAG22 ANTIGEN.
74354 MW; F6953C3DBE8E6AC8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301 QGGYGGL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31.1%; Score 839.5; DB 1; 37.4%; Pred. No. 2.5e-32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22; Mismatches 210;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Antigen; Repeat; Signal; Complete proteome.
SIGNAL 1 30 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AE007040; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  = - :-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TIGR; MT1807; --
Tubercuclist; Rv1759c; --
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best_Local Similarity 37.4%
Matches 249; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; Z95890; CAB09322.1;
Peterson J., DeBoy R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE 914 AA;
                                                                                                                                                                                                                                                                          SUBFAMILY.
```

us-09-490-291-8.rsp

g

g ò

õ g

```
T00211; -.
IPR004186; EBNA1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mycobacterium tuberculosis.
                                                                                                                                                                                                                           Best_Local Similarity 46.0% Matches 209; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                          641 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                         FRANSFAC; T00211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-1773;
                                                                                                                                   87
                                                                                                            3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-H37Rv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Y747_MYCTU
053810;
                                                  InterPro;
                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                      Query Match
                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              201
      SOR KE DR.
                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lib.sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=84270667; PubMed=6087149;
Baer R., Bankler A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
Tuffnell P.S., Barrell B.G.;
"DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
Nature 310:207-211(1984).
                                           436 ----GAAAAVGAG-----QGGYGGVGSGASAASALSSPQASSRVSSAVSNLVASGP 486
                                                                                          782
666 GAAGLFANGGAGGAGGTGSTAGGAGGAGGAKAHGGTGGPGGNGGSTGAGGTGGAGGP 725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOMEWHAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 1-26 FROM N.A.
MEDLINE-86259739; PubMed-3460083;
MEDLINE-86259739; PubMed-3460083;
Sample J., Hummel M., Braun D., Blrkenbach M., Kieff E.;
"Nuclectide sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: a probable transcriptional initiation site.";
Proc. Natl. Acad. Sci. U.S.A. 83:5096-5100(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FUNCTION: INVOLVED IN LATENT CYCLE. EBNA-1 FUNCTIONS IN THE MAINTENANCE REPLICATION OF EBV EFISOME. TRANSACTIVATING FACTOR FOR THE ORIGIN AND ENHANCER FUNCTIONS OF ORIP.
SUBCELLULAR LOCATION: NUCLEAR. FREE IN THE NUCLEOPLASM, SOMEWHAA ASSOCIATED WITH THE CHROMATIN AND HARDLY, IF AT ALL ASSOCIATED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBCELLULAR LOCATION.
MEDIINE-90266473; PubMed-2161150;
MEDLINE-90266473; PubMed-2161150;
MEDLINE-Sample C., Kieff E.;
"Subnuclear localization and phosphorylation of Epstein-Barr virus latent infection nuclear proteins.";
Virology 176:563-574(1990).
                                                                     X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 469-607.

MEDLINE-96006523; PubMed=7553871;

MEDLINE-94006523; PubMed=7553871;

MEDLINE-94006523; PubMed=7553871;

BOCHWART A. M., Frapplir L.;

Edwards A.M., Frapplir L.;

ECTYSTAL STRUCTURE Of the DNA-binding domain of the Epstein-Barr
                                                                                                                                                                                                                                                                                                                                                                                                                                             Epstein-barr virus (strain B95-8) (Human herpesvirus 4).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Gammaherpesvirinae; Lymphocryptovirus.
                                                                                                                                                                                                                                                                                                                                21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  virus origin-binding protein EBNA 1.";
Cell 83:39-46(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; V01555; CAA24816.1; -. EMBL; M13941; AAA45889.1; -. PIR; A03773; QQBE31. PIR; S33021; S33021.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WITH THE NUCLEAR MATRIX.
                                                                                                                                                                                                                                                                                                                                                                                                   EBNA-1 nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=10377;
                                                                                                                               487 TNSAAL 492
                                                                                                                                                                             783 GGSAGL 788
                                                                                                                                                                                                                                                                                         EBN1_EBV
                                                                                                                                                                                                                                             RESULT
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ۲.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100 AGGVGQG------GLGGQ--GAGQGAGAAAAAAAGGAGQGGYGGLGSQGA 140
                                                                                                                                                                                                                                                                                                                                                                                                 57 PGGSGSGPRHRDGVRRPQKRPSCIGCKGTHGGTGAGAGGAGGAGGAGGGAGAGGGAG 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      141 GRGGSGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAAGGAGQGGYG 200
                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                            26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          s.
                                                                                                                                                                                                                                                          40 GAGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAA 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Mycobacterium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Harris
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-9829597; PubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris Gordon S.V., Elgimeler K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Bovles R., Deviln K., Feltwell T., Gentles S., Hamlin N., Holroyd (Hornsby T., Jagels K., Kroph A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
Pfam; PF02905; EBNA1; 1.
Nuclear protein; DNA-binding; Transcription regulation; Activator;
                                                                                                                                                                                                                                                                                       261. QGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGQGGGYGGLGSQGAGRGGLGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAGGAGQGGLGGQGAGAAAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        381 GVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAGAAAA
                                                                                                                                                                                                             75;
                                                                                                                                                           Length 641;
                                                                                                                                                                                                           Indels
                                                                               GLY/ALA-RICH.
4D161653E16FC341 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical PE-PGRS family protein Rv0747 precursor.
RV0747 OR MT0772.5 OR MTV041.21.
                                                                                                                                                      30.7%; Score 828.5; DB 1;
46.0%; Pred. No. 6.4e-32;
tive 18; Mismatches 152;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           801 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           441 VGAGQG-GYGGVGSGASAASAASRLSSPQASSR 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                               352 G
56427 MW;
```

```
554
                                                                                                                                           350
                                                                                                                                                                                                                                                               672
                                                                                                                                                                                                                                                                                                                                           732
                                                                                                                                                                                                                                                                                                      459
                                                               296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 10
                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HID DATE OF THE SECOND OF THE 
                                                               ò
                                                                                                    g
                                                                                                                                       ò
                                                                                                                                                                              g
                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce.or send an email to licenseelsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25;
                                                 STRAIN-CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bishai W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GNGGAGGSGA-PGAGAGGNGGAGGLFGSGGAGGASTDVAGGAGGAGGAGGNAGMLFGAAG 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QGGYGGLGSQGA--GRGGLGGQGA--GAAAAAGGVGQGGL--GGQGAGQGAGAAAAAAGG 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VGGVGGFSNGGATGGAGGAGGAFGAGRERGSGGSGNLTGCAGGAGGNAGTLATGDGG 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GNGGNGGASTGGGDGGPGGAGGTGVLIGNGGNGGSGGTGATLGKAGIGGTGGVLLGLDGF 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IGGQGA-----GAAAAAAGGAGQGGYGGIGSQGAGRGGQGAGAAAAAAGGAGQGGYG 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GYGGLGGQGAGQGAGAAAAAAAGGA-GQGGYGGLGSQ---GAGRGGQGAGAAAAAAGGAG 72
                                                                                                                                                                            "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AG-----QGGYGGLGSQGA-----GRGGSGGQG------AGAAAAAG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       256 AGGTGGASRSGGFGGAGGAGGDAGMFFGSGGSGGAGGISKSVGDSAAGGAGGAPGLIGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------GAGQGGYGGLGSQG----AGRGGLGGQGA-----GAGQGGYGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 376 TAPASTSPLHTLOODVINMVNDPFOTLTGRPLIGNGANGTPGTGADGGAGGWLFGNGGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----QGGYGGLGGQGA---GQGGYGGLGSQGA-GRGGLGGQGAGAAAAAGGAGQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYPOTHETICAL PE-PGRS FAMILY PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 165;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----GAAAAAAGGAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 801;
                                                                                                                                                                                               laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EA54C9BF45A00F41 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .9e-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S (IN REF. 2
G (IN REF. 2
G (IN REF. 2
G (IN REF. 2
I (IN REF. 2
P (IN REF. 2
A (IN REF. 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30.4%; Score 819.5; F
36.3%; Pred. No. 1.9e-
ive 19; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           G - V G (1)

K - V G (1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RV0747.
G -> S
R -> G
R -> G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TIGR; MT0772.5; -..
TubercuList; Rv0747; -.
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE; I.
Hypothetical protein; Repeat; S:
SIGNAL 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AL021958; CAA17514.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AE006968; AAK45011.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65407
Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     295
300
338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       188
225
227
295
330
338
377
577
580
801 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31
                                                                                                                                                                                                                                                           SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Sim.
Matches 242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            184
õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ilensib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                          402
                                                                                                                                                                                                                                                                                                                                                      671
                                                                                                                                                                                                                                                                                                                                                                                                                                   458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ۵.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacterium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-9682987; PubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harriss Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harriss Gordon S.V., Elgimeler K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd Elbrosby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; "Deciphering the blology of Mycobacterium tuberculosis from the Complete genome sequence.";
614 AAGVSGGAGGAGGDAGLLSLGASGGAGGSGGSSLTAAGVV--GGIGGAGGLLFGSGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                               403 GAAAAAAGGAGGGTGGLGGQGVGAGGLGGQGAGAAAAVGAGQGGYGG----VGSGASAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SGEFSNSGNGGAGGAGGAGGLLVGSGGAGGAGAATGGDGGAGGKSGAFGLGGDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SAAASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCDVLIQ
                                                                                                                GQGAGQGGYGGLGSQGA--GRGGLGGQGAGAAAAAAAAAAGGGGGL-----GGQGAGAAAA
                                                                                                                                                                                         GSPAGAGGIGGAGGNGGLFGAGGTGGAGGGSTLAGGAGGAGGNGGLFGAGGTGGAGSHST
                                                                                                                                                                                                                                                                      AAG-GAGQGGLGGQ------GAGQGAGAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AGGATGLSGAFHIGGKGGVGGSAVLIGNGGNGGNGGNSGNAGKSGGA---PGPSGAGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-MAY-2000 (Rel. 39, Last sequence update)
16-0cr-2001 (Rel. 40, Last aminotation update)
Hypothetical PE-PGRS family protein Rv0278c precursor.
RV0278C OR MT0291 OR MTV035.06C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    957 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30-MAY-2000 (Rel. 39, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              790 LLLGEN 795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          519 ALLGHH 524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Y278_MYCTU
P56877;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bishai W.;
```

σ

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                     22;
                                                                                                                                                                                                                                                                                                                                            276
                                                                                                                                                                                                                                                                                           200
                                                                                                                                                                                                                                                                                                           227
                                                                                                                                                                                                                                                                                                                            560
                                                                                                                                                                                                                                                                                                                                                                                              680
                                                                                                                                                                                                                                                                                                                                                                                                               364
                                                                                                                                                                                                                                                                                                                                                                                                                                                413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 860
                                                                                                                                                                                     265 GVGGSGGAGSSGGAGGAGGLFGAGGTGGHGGFADSSFG-GVGGAGGAGGLFGAGGEG 323
                                                                                                                                                                                                       GYGGLGSQGAGRGGLGGQGA----GAAAAAGGVG-----QGGLGGQGAGQGAGAAAA 123
                                                                                                                                                                                                                ----QGAGAAAAAGGAGQGGYGQQGGYGGLGSQGA-----GRGGLGGQGAGAAA 326
                                                                                                                                                                                                                                                                                                                                                                                                                             681 PAGYGGAGGNGGLFANGGAGGAGGFNAAGGNGGNGGLFGTGGTGGAGTNFGAGGNGGNGG 740
                                                                                                                                  ; Score 807; DB 1; Length 957;
; Pred. No. 8e-31;
21; Mismatches 195; Indels 174; Gaps
                                                                                                                                                                                                                                                 17 GYGGLGGQGAGQGAGAAAAAA--GGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQG
                                                                                                                                                                                                                                        124 AGGAGQGGYGGLGSQGAGRGGSGGO----GAGAAAAAGGAGQGGYGGLGSQGA----
                                                                                                                                                                                                                                                                         174 GRGCLGGQGAGAAAAAAAGG-----AGQGGYGGLGGQGAGQGGYGGL-----
                                                                                                                                                                                                                                                                                                           ------GSOGAGR-GGLGG-----
                                                                                                                                                                                                                                                                                                                           501 APASTSPLHTLQQNVLNVVNEPFQTLTGRPLIGNGANGTPGTGADGGAGGWLFGNGANGT
                                                                                                                                                                                                                                                                                                                                           228 QGAGAAAAAGG--AGQGGLGGQGA-GAAAAAGGA------GQGGYGGLGSQGAGRGG
                                                                                                                                                                                                                                                                                                                                                      621 IGGAGGAGGVSLLIGSGGTGGNGGNSIGVAGIGGAGGRGGDAGLLFGAAGTGGHGAAGGV
                                                                                                                                                                                                                                                                                                                                                                                                                                               365 GQGAGAAAAAAAAAGGVRQGGYG-----GLGSQGAGRGGQGAGAAAAAAAGGAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------VGAGGLGGQGAGAAAVGAGQG
                                                                         HYPOTHETICAL PE-PGRS FAMILY PROTEIN
                         Tuberculist; Rv0278c; -.
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE; 1.
Hypothetical protein; Repeat; Signal; Complete proteome.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TD25_MYCTU STANDARD; PRT; 603 AA. Q10637; 01-0CT-1996 (Rel. 34, Last sequence update) 16-0CT-2001 (Rel. 40, Last annotation update) Hypothetical PE-PGRS family protein Rv1325c precursor.
                                                                                         M -> I (IN REF. 2).
MISSING (IN REF. 2).
R -> G (IN REF. 2).
71EBABD417FBA47C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                               AAAAAGGAGQGGL----GGQGAGAAAAAAGGAGQGGL-----
                                                                                                                                                                                =
=
=
                                                                                 RV0278C.
      EMBL; AL021930; CAA17353.1; -.
EMBL; AE006936; AAK44511.1; ALT_INIT
TIGR; MT0291; -.
                                                                                                                   81905 MW;
                                                                                                                                    29.9%;
36.0%;
                                                                                                                                   Query Match
Best Local Similarity 36.0%
Matches 219; Conservative
                                                                 30
                                                                                        40
163
807
                                                                                         40
158 1
807 8
957 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 447 GYGGVGSGA 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GTGGDGGHA
                                                                                        CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                          CHAIN
                                                                                                                                                                                                      75
                                                                                                                                                                                                                                                                                                                                                                             277
                                                                                                                                                                                                                                                                                                                                                                                                               327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 861
                                                                                                                                                                                                                                                                                                          216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eddag
Q
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                      ð
                                                                                                                                                                                     셤
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                       οŽ
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license&ib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
                                                                                                                                                                                                                                                                  : ۵
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LFGNGGAGGAGGDGPGNGAAGGAGGAGGLLFGSGGAGGPGGVGNTGTGGLG---GDGGAA 218
                                                                                                                                                                                                                                                                                                                                                   S:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   70 GAGQGGYGGLGSQGAGRGGLGGQ------GAGAAAAGGVGQGGLGGQGAGQGAGAAA 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122 AAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLGGQ 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        113 LGRPLI-GNGANGAPGTGQ-----AGGAGGLLYGNGGAGGSGAPGQ-AGGPGGAAG 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10 MGRIRIRGYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGQQGAGAAAAAG 69
                    Mycobacterium tuberculosis.
Bacteria, Actinobacteria, Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.

HYPOTHETICAL PE-PGRS FAMILY PROTEIN
RY1325C.

GLY-RICH.

G -> V (IN REF. 2).

G -> D (IN REF. 2).

G -> D (IN REF. 2).

L -> F (IN REF. 2).

L -> F (IN REF. 2).
                                                                                                                                                                                                                      MEDLINE-9892987; PubMed-9634230;

Gole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris Gordon S.V., Elgimeler K., Gas S., Barry C.E. III, Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd Hornsby T., Jagels K., Krooph A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quall M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.B., Taylor K., Miltchead S., Barrell B.G.;

"Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 603;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Repeat; Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29.0%; Score 782.5; DB 1;
43.0%; Pred. No. 7.7e-30;
ive 13; Mismatches 202;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AE007010; AAK45630.1; ALT_INIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Tuberculist; Rv1325c; -.
InterPro; IPR000084; PE_region.
Pfam; PF00934; PE; 1.
ProDom; PD001223; PE_region; 1.
Hypothetical protein; Repeat; Sign SIGNAL
RV1325C OR MT1367 OR MTCY130.10C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             603
132
135
6
135
6
337
508
L
49575 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
STRAIN-CDC 1551 / Oshkosh;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; Z73902; CAA98089.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                114
132
135
337
508
603 AA;
                                                                                                                                           [1]
SEQUENCE FROM N.A.
STRAIN-H37Rv;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Sim
Matches 222;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NACUTE 3.53:1/F181(1200).

-1 - FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).

-1 - SUBCELLULAR LOCATION: Cell wall (Potential).

-1 - MISCELLANBOUS: THIS PROTEIN CONTAINS 67% GLYCINE RESIDUES.

-1 - MISCELLANBOUS: PROTEIN RESIDUES ARE CAPABLE OF FORMING A BETA-PLEATE SHEET COMPOSED OF 8 ANTI-PARALLEL STRANDS.

-1 - MISCELLANBOUS: THE GLYCINE REGION IS COMPRISED OF TWO RELATED FAMILIES OF REPEATS, F1 AND F2, EACH REPEAT CONTAINING ABOUT 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
Condit C.M., Meagher R.B.;
"A gene encoding a novel glycine-rich structural protein of petunia.";
Nature 323:178-181(1986).
                                                                                                                                                 440
                                                      GAGAA--AAAAAGGAG-----QGGYGGLGGQG---AGCGGCYGGLGSQG-----AGRG 223
                                                                                                                                                                                                         327
                                                                                                                                                                                                                                         453
                                                                                                                                                                                                                                                                                                                                                                          513 GAGGAGGLGIAGDGGNGGKG-----GRAGMVGNGGDGGAGGASVVANGGVGGSGGNATLI 567
224 GLGGQGAGAAAAAGGAGQGGLGGQGAGAAAAAAGGAGQGGYGGLGSQGA-----GRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Petunia hybrida (Petunia).
Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;
Permatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
NCBI_TaxID-4102;
                                                                                        275 GGGGGLFSNGGAGGAGGFGVSGSAGGNGGTGGDGGIFTGNGGTGGTGGTGTGNQLVGGEG
                                                                                                                                                                                                  GQGAGAAAAAAGG-AGQGGYGGQG--AGQGGYGGLGSQ----GAGRGGLGGQGAGAA-AA
                                                                                                                                                                                                                       454 AGGAGGTGIEFGSVGGAGGAG-GNAAGLSGAGGAGGFGETAGDGGAGGNAGLLNGDG
                                                                                                                                                                                                                                                                        383 RQGGYGGLGSQG-AGRGGQGAGAAAAAGGAGQGGTGGLGGQGVGA-GGLGGQGAGAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLYCINE-RICH CELL WALL STRUCTURAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAR-1989 (Rel. 10, Created)
01-MAR-1989 (Rel. 10, Last sequence update)
01-AUG-1992 (Rel. 23, Last annotation update)
Glycine-rich cell wall structural protein 1 precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLY-RICH,
C8541C549417D18C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A26099; A26099.
Cell wall; Structural_protein; Repeat; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                            GNGGNGGNGSAPGKGGAGGTAGLLGLNGSPGLS 603
                                                                                                                                                                                                                                                                                                                                                                                                                     441 VGAGGGGGGGGGS--GASAASAASRL---SSPQAS 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    384 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   384 G
28777 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; X04335; CAA27866.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   384 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GRP1_PETHY P09789;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNAL
                                                                                                                                                                                                  276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                            ŝ
                                                                                                                                                              g
                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                          ç
                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lsb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                     22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CALTOMOSOME / ...

-I- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND
NUCHAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.

-I- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER
INTO AN EXTERNIBLE 3D EXPRONCE.

-I- SUBCELLULAR LOCATION: EXPRACELLULAR MATRIX OF ELASTIC FIBERS.

-I- PTW: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.
                                                                                                                                          97 AAAAGGVGQGGLGGQGAGQAGAAAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAA 156
                                                                                                                                                                                                                                                                                                                                                                                                        285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   144 GGGAGSGGGFGAG-----GGVGG-GAGA------GGGVGGGGGFGG-----GGGGGVG 184
                                                                                                            93
                                                                      40 GAGQGGYGGLGSQGAGRG---GQGAGAAAAAGGAGQGGYGGLGSQGAGRGGLGGQCAGA 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ဌ
                                                                                            QGAGAAAAAAGGAGQGGYGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-BALB/C; TISSUE-Lung;
MEDLINE=95130069; PubMed=7829060;
Wydner K.S., Sechler J.L., Boyd C.D., Passmore H.C.;
"Use of an intron polymorphism to localize the tropoelastin gene t
mouse chromosome 5 in a region of linkage conservation with human
                                                                                                                                                                                                                                                                                        217 SQGAGRGGLGGQGAGAAAAAGGAGQGGLGGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGG
                                                                                                                                                                      94 -GAGGGGGLGGGGGGAGGGFG------GGAGGGAGGGLGG-GGGLGGGGGGGGG
                                                                                                                                                                                                                  AGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                       185 G-GSGHGGGFGAGGGVGGAGGGLGGGVGG------GGGGGSGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           286 GAAGGVGGGGFGGGGGGVGGGSGHGGGFGAG-----GGVGGGAGGGLG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             397 RGGQGAGAAAAAGGAGQGGTGGLGGQGVGAG-GLG-GQGAGAAAAVGAGQG-GYGGVGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muridae; Musinae; Mus
                                       84;
 Length 384;
                                     Indels
                                    10; Mismatches 130;
27.3%; Score 736; DB 1; 46.8%; Pred. No. 7.5e-28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    860 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Elastin precursor (Tropoelastin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata;
Mammalla; Eutheria; Rodentia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
CTRAIN-BALB/C; TISSUE-Lung;
                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mus musculus (Mouse).
                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-10090;
                 Local Simi
hes 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     G 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       G 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ELS_MOUSE
P54320:
 Query Match
                  Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                       337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ELS_MOUSE
                                                                                                                                            à
                                                                                                                                                                              g
                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                     25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last annotation update)
Glycine-rich cell wall structural protein precursor.
Arabidopsis thallana (Mouse-ear cress).
Bukaryota; Viridiplanane; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; wagnollophyta; eudioctyledons; core eudicots; Rosidae; eurosids II; Erassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Quigley F., Villiot M.L., Mache R.;
Quigley F., Villiot M.L., Mache R.;
"Nucleotide sequence and expression of a novel glycine-rich protein
gene from Arabidopsis thaliana.";
Plant Mol. Biol. 17:349-952(1991).
-1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                         472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           646
                                                                                                                                                                                                                                                                                                                                                                                                                                    532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GAGTPAAAAAAKAAAKAGAAKYGAAGGLVPGGPGVRLPGAGIPGVGGIPGVGGIPGVGGPGI 412
                                                                                                                                                                                                                                                                                                                                          151
                                                                                                                                                                                                                                                                                                                                                                                                    152 -AAAAAAGGAGQGGYGGLGSQGA------GRGGLGGQG--AGAAAAA 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GAGAGVPGFGAGAVPGSLAASKAAKYGAAGGLGGPGGLGGPG------GLGGPGGLGG 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     362 QG-----AGQGAGAAAAAAAAAGGVRQ---GGYGGLGSQGAGRGGQGAGAAAAAAGGAG 413
                                                                                                                                                                                                                                                                             GAG---AAAAAAGGAGQGGYGGLGS-----QGAGRGGLGG-QCAGAAAAAGGVGQ 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            533 AKAAAKAGLGPGVGGVPGGVGGIPG----GVGVGGVPG-GVGPGGVTGIGAGPGGLGG 587
                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                            QGAGAAA-AAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAG-GAGQGGYGGQGGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         588 AGSPAAAKSAAKAAAKAQYRAAAGLGAGVPGFGAGAGVPGFGAGAGVPGFGA-GAGVPGF
                                                                                                                                                                                                                                                                                                                                          GG----LGGQGAGQGAGAAAAAGGAGQGGYGGLG--SQGAGRGGSGGQGAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                AGGAGQGGYG-GLGG-QGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAG-GAGQGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        G-GLGSQCAGRGGLGGQGAGAAAA--AAAAGGAGQGGLGGQGAGAAAAAAGGAGQGGLGG
                                                                                                                                                                                     103;
                                                                                                                                                     DB 1; Length 860;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                414 QGGTGGLGGQGVGAGGLGGQGAGAAAVGAGQGGYGGV-----GSGASA 457
                                                                                                                                                   27.0%; Score 729.5; DB 1; Length (42.6%; Pred. No. 2.5e-27; ive 23; Mismatches 179; Indels
                                                                                                                                                                                                                  15 IRGYGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLG-SQGAGRGG-----
                                                                                                                        OCOBESAAELEDD7F1 CRC64;
                                                                      Structural protein; Repeat; Signal; Connective tissue. SIGNAL 1 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          338 AA
                                                                                                         ELASTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                           IPR003979; tropoelastin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-cv. Columbia;
MEDLINE-92003708; PubMed-1912511;
                                                                                                                      71955 MW;
                                                           PRINTS; PR01500; TROPOELASTIN
              EMBL; U08210; AAA80155.1;
                                                                                                                                                                                   226; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                         860
                                                                                                                     860 AA;
                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                eurosids II; Bra
NCBI_TaxID=3702;
                           MGD; MGI:95317;
InterPro; IPR003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GRP_ARATH P27483;
                                                                                                                        SEQUENCE
                                                                                                                                                      Query Match
                                                                                                                                                                       Local
                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                         106
                                                                                                                                                                                                                                                                                                                                                                      413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     647
STTXRDRAFTS
                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                  ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is neor removed. Usage by. and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ilential.).
                                                                                                                                                                                                                                                                                                                          23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGGFGGGAGGGLGGGAGGGTG---GGFGG----GAGGGAGGGAGGG-----FGGGAGGG 263
                                                                                                                                                                                                                                                                                                                                                                                                                 99 AAGGVGQGGLGGQGAGAAAAAAAAGGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAAG 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                      39 GGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAA 98
                                                                                                                                                                                                                                                                                                                                                                        Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Mammalia; Eutheria; Rodentia; Sciurognathl; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                             --GGAG-GGLGG-GAGGGLG-----GGHG-GGIGG----GAGGGAGGGLGGGHGGGIGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GAGRGGLGGQGAGAAAAAGGAGQGGLGGQ-GAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GAGAAAAAAAGGGGYGGYGGQGAGQGGYGGLGSQGAGGGGGGGGGGGAGAAAAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                          77;
                                                                                                                                                                                               POTENTIAL, GELL WALL STRUCTURAL GLYCINE-RICH CELL WALL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE-90275605; PubMed-2190691;
Mehrel T., Hohl D., Rothnagel J.A., Longley M.A., Bundman D.,
Cheng C., Lichti U., Bisher M.E., Steven A.C., Steinart P.M.,
Yuspa S.H., Roop D.R.;
"Identification of a major keratinocyte cell envelope protein,
loricrin.";
                                                                                                                                                                                                                                                                                        26.5%; Score 714.5; DB 1; Length 338; 49.0%; Pred. No. 6.6e-27; Live 6; Mismatches 117; Indels 77;
                                                                                                                                                                                                                                            21 338 GLY-RICH.
338 AA; 23891 MW; 046A6E8C1A4E89EB CRC64;
SUBCELLULAR LOCATION: Cell wall (Potential).
                                                                                                                                              EMBL; X58338; CAA41249.1; -. PIR; S17732; KNMU. Cell wall; Structural protein; Repeat; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  481 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GCQCAGAAAAAGGAGQGGTGGLGGQGVGAGG 429
                                                                                                                                                                                                                               PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Rel. 16, (Rel. 16, (Rel. 41,
                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1990 (
01-NOV-1990 (
15-JUN-2002 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LORI_MOUSE
P18165;
                                                                                                                                                                                                                                                                                                           Best Local Sim
Matches 192;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Loricrin.
                                                                                                                                                                                                                                            DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LORI_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                             .92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      306
 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               G-GGSSCGGGSSGG-GSSCGGSSGGGGSSGGGGGXSSGGGGXSSQQTSQ 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GRGGLGG----QGAGAAAAAGGVGQGGLGGQGAG---QGAGAAAAAAGG-----AG 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64 GGGGSGGCVKYSGGGGGSSCGGGYSGGGGGSSCGGGSSCGGGSSCGGGYSGGGGSSCG 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QGGYGGLGSQGAGRGGSGG-----QGAGAAAAAGGAGQGGYGGLGSQGAGRGG----LG 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   239 AGGGGGGGGAGAAAAAAGGAGGG-----GYGGLGSQGAGRGGQGAGAAAAAAGGAGGG 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       244 GGYSGGGGTSCGGGSSGGGGGSSQQYQCQSYGGGSSGGSSCGGGYSGGGGSSCGGGYSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GYGGQGAGQGGYGGLGSQ--GAGRGGLGGQGAGAAAAAAAAGGAGQGGLGG------Q 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15 GCGKTSG------GGGGGGGYSGGGSGGGGGSSGG---SSCGGGGGGSYGG-GSSCG 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25 GAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQGA 84
                                                                                                     356 TSCAPQQSYGGGSSGGGGSCGGGSGGGGCCXSGGGGSSGGGG-GYSGGG-GGCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GOGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY-GGLGSQGAGRGGLGGQGAGAAAAAGG
                                          STRAIN-BALB/c;
MEDLINE-95256248; PubMed-7738016;
DiSepio_D., Jones A., Longley M.A., Bundman D., Rothnagel J.A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             65;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 25.8%; Score 697.5; DB 1; Length 481; Best Local Similarity 37.7%; Pred. No. 5e-26; Matches 178; Conservative 35; Mismatches 194; Indels 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GAAAAAAGGGGTGGLGGQGVGAGGLGGQGAGAAAVGAGQGGYGGVGSG 454
                                                                                                                                                                                                                                                                                                                                                                                                                                              481 AA; 37830 MW; 97349A786FF239FE CRC64;
                                                                                                                                                                                                                                                                                                                                                  EMBL; M34398; AAA3944.1; -. EMBL; U09189; AAA82152.1; -. PIR; A35628, A35628. HSSP; P02876; 9WGA.
Sell 61:1103-1112(1990).
                                                                                                                                                                                                                                                                                                                                                                                                              MGD; MGI:96816; Lor
                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                Keratinization
                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                414
 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
```

Search completed: December 18, 2002, 16:18:22 Job time : 15:1044 secs

```
December 18, 2002, 16:13:04; Search time 30.9992 Seconds (without alignments) 3509.535 Million cell updates/sec
                                                                                                                                                                                                          US-09-490-291-8
2700
1 MASWIGGQQMGRIRIRGYGG......GLSGCDVLIQALLGHHHHH 528
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                    - protein search, using sw model
                                                                                                                                                                                                              Title:
Perfect score:
                                                                                    OM protein
                                                                                                                                                                                                                                                      Sequence:
                                                                                                                            Run on:
```

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

671580 seqs, 206047115 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_manmal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:* sp_v1rus:*
sp_vertebrate:*
sp_unclassified:* sp_rvirus:*
sp_bacteriap:* sp_archea:* sp_bacteria:* sp_rodent:* sp_plant:* SPTREMBL_21:* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

sp_archeap:*

08wsw4 nephila cla 09biu7 argiope tri 09biu3 dolomedes t 017434 nephila cla 09biu4 dolomedes t 09biu1 argiope aur 06897 araneus dia 09bit5 nephila mad 09bit7 nephila mad 09biu0 latrodectus 09biu0 latrodectus 09biu0 atrodectus 046171 nephila cla 046172 nephila cla Description SUMMARIES OSIZ 08 08 08 046172 Q8WSW4 Q9BIU7 Q9BIU3 017434 Q9BIU4 Q9BIV1 Q9BIT5 Q9BIT7 0964F4 09NHW2 16 DB 563 1953 360 2639 1715 1884 Query Match Length Score 2313 2290 1570.5 1405.5 1305.5 1305.5 1009 1042.5 1003 942.5 942.5 Result 44 11 11 12 13 14 14 16 17 . 2

053557 mycobacteri 084199 mycobacteri 044358 nephila cla	O53559 mycobacter1 Q9bit4 nephila sen	Ognhw4 nephila cla	Q8V1Z1 Mycobacteri O53552 mycobacteri	002402 pinctada fu	Q9blu6 arglope tri	Q9bit2 plectreurys	053435 mycobacter1	Q9gub5 galler1a me	Q9biu8 argiope tri	053215 mycobacter1	Q9bit8 latrodectus	044359 nephila cla	O53439 mycobacter1	Q8vk17 mycobacter1	053845 mycobacter1.			O9bit6 nephila mad	O53844 mycobacter1	Q9nhwl nephila mad	006810 mycobacter1		053395 mycobacter1	.08vj23 mycobacter1
053557 Q8VIY9 044358	053559 Q9BIT4	O9NHW4	053552	002402	901860	Q9BIT2	053435	Q9GUB5	Q9BIU8	053215	Q9BIT8	044359	053439	Q8VK17	053845	Q8VKD2	09BIU9	Q9BIT6	053844	O9NHW1	006810	053556	053395	Q8VJ23
16 16 5	16 5	2 4	10		S		16		S	16	ស	'n	16	16	16	19	S	S	16	S	16	19	16	16
1079 1217 871	1489 253	2249	1381	738	444	912	167	1468	1.002	1665	399	907	853	1408	882	879	651	233	749	979	1329	714	1538	1507
34.3 34.3 34.1	33.8 33.4	33.3	33.0	33.0	32.9	32.6	32.6	32.5	32.3	32.2	32.1	31.9	31.8	31.8	31.8	31.7	31.6	31.5	31.5	31.4	31.4	31.2	31.0	30.9
927 926 921	912.5	900	892	890.5	888.5	881	880	877.5	873	868.5	866.5	862	828	828	857.5	857	852	850.5	850	847	847	843.5	836	835.5
17 18 19	20 21	22	24	25	56	27	28	53	30	. 31	32	33	34	35	36	37	38	33	40	41	42	43	44	45

## ALIGNMENTS

```
Beckwitt R., Arcidiacono S., Stote R.; "Evolution of repetitive proteins: spider silks from Nephila clavipes (Tetragnathidae) and Araneus bicentenarius (Araneidae)."; Insect Blochem. Mol. Biol. 28:121-130(1998).
                                                                         Spidroin 1 (Fragment).
Nephila clavipes (Orb spider).
Nephila clavipes (Orb spider).
Nephila clavipes (Arthropoda, Chelicerata; Arachnida; Araneae; Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
                                                                                                                                                                                                                                                              [2]
SEQUENCE OF 449-544 FROM N.A.
MEDILINE-99148667; PubMed-9487707,
Arcidiacono S., Mello C., Kaplan D., Cheley S., Bayley H.;
Purification and characterization of recombinant spider silk
expressed in Escherichia coli.";
Appl. Microbiol. Blotechnol. 49:31-38(1998).
                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-90384959; Pubmed-2402494;
Xu M., Lewis R.V.;
"Structure of a protein superfiber: spider dragline silk.";
Proc. Natl. Acad. Sci. U.S.A. 87:7120-7124(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
Beckwitt R.;
Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
EMBL; U20329; AAC38957.1; -.
HSSP; P10969; 1 WGT.
NON_TER
                              01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 544 AA.
   PRT;
   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                             [3]
SEQUENCE FROM N.A.
MEDLINE=98318805;
046171
```

~

<u>,,</u>

49410 MW; 1EF0CE9269A832E2 CRC64;

617 AA;

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                 =:1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    O8WSW4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Basel
                                                                                                                                                                                                                                                                                                                              421
                                                                                                                                                                                                  266
                                                                                                                                                                                                                                      326
                                                                                                                                                                                                                                                                                                                                                400
                                                                                                                                                                                                                                                                                                                                                                481
                                                                                                                                                                                                                                                                                                                                                                                   460
                                                                                                                                                                                                                                                                                                                                                                                                                     520
                                                                                                                                                                                                                                                                         333
                                                                                                                                                                                                                                                                                                            365
                                                                                                                                                                                                                                                                                                                                                                                                                                      601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
Q8WSW4
 S
                                                                        a
                                                                                                           g
                                                                                                                                             a
                                                                                                                                                                                셤
                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                            g
                                                                                          ð
                                                                                                                            á
                                                                                                                                                                à
                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                        å
                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                ç
                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-94165058; PubMed-8120021; MEDLINE-94165058; PubMed-8120021; Beckwitt R., Arcidiacono S.; Sequence conservation in the C-terminal region of spider silk proteins (Spidroin) from Nephila clavipes (Tetragnathidae) and Araneus bicentenarius (Araneidae)."; J. Biol. Chem. 269:6661-6663(1994).
                                                                                                                                                                256
                                                                                                                                                                                240
                                                                                                                                                                                                                    300
                                                                                                                                                                                                                                     376
                                                                                                                                                                                                                                                                        436
                                                                                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-98318805;
Beckwitt R., Arcidiacono S., Stote R.;
Beckwitt R., Arcidiacono S., Stote R.;
Brownition of repetitive proteins: spider silks from Nephila clavipes (Tetragnathidae) and Araneus bicentenarius (Araneidae).";
Insect Blochem. Mol. Biol. 28:121-130(1998).
BMBL: 037520; AAC04504.1; -.
NON_TER 617 617
                                                                                                                                   316
                                                                                                                                                                                                                                              SQGAGRGGSGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQ 196
                                                                                                                                                                                                                                                                                                            496
                                                                                       GGLGSQGAGRGGLGGQGAGAAAAAGGVGQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLG 136
                                    Gaps
                                                    9/
                                                             GYGGIGGQGAGQGAGAAAAAAGGAGQGGYGGIGSQGAGRGGQGAGAAAAAAGGAGQGGY
                                                                                                AAAAVGAGGGGGGGGGGGASAASAASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTI
                                                                                                                                                              GGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAGGAGQGGLGGQGAGAAAAA
                                                                                                                                                                                                 GGAGQGGYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGQGAGQGGYGGLGSQGAGRGG
                                                                                                                                                                                                                                    AAAGGVRQGGYGGLGSQGAGRCGQGAGAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Araneae;
Nephila.
                                    ö
                  Length 544;
                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Dragline slik protein spidroln 1 (Fragment).
Nephila clavipes (Orb spider).
Bukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida;
Araneomorphae; Entelegynae; Arancoidea; Tetragnathidae;
                Score 2555; DB 5; 1
Pred. No. 1e-122;
0; Mismatches 2;
CC611B3551945615
                                                                                                                                                                                                                                                                                                                                                                                                            617 AA
                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                            SNVVSQIGASNPGLSGCDVLIQALL 521
44107 MW;
                 94.68;
                         Best Local Similarity 99.6
Matches 503; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
544 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
SEQUENCE
                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                          046172 046172;
                                                    17
                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                          437
                                                                                                                                                                                                                                                                                                                                                              481
                                                                                       11
                                                                                                        19
                                                                                                                          137
                                                                                                                                            121
                                                                                                                                                              197
                                                                                                                                                                                                 257
                                                                                                                                                                                                                  241
                                                                                                                                                                                                                                    317
                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                       377
                                                                                                                                                                                                                                                                                        361
                                                                                                                                                                                                                                                                                                                           421
                                                                                                                                                                                                                                                                                                                                             497
                                                                                                                                                                                                                                                                                                                                                                                         RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                  တ္တ
                                                    à
                                                                     셤
                                                                                       õ
                                                                                                      q
                                                                                                                          à
                                                                                                                                          g
                                                                                                                                                             à
                                                                                                                                                                            g
                                                                                                                                                                                                ò
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                            8
```

```
211
                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                         265
                                                                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                   300
                                                                                                                                                                                                                                                                                                                                                                                  332
                                                                                                                                                                                                                                                                                                                                                                                                            420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                            94 AG--AAAAAGGVGQGGLGGQGAGAGAAAAAAGAGQGGYGGLGSQGAGRGGSGGQGAG 151
                                                                                                                                                                                                                                                                                                                    GLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGQGAGGGGYGGLGSQGAGRGGLGGQGAGAA 325
                                                                                                                                                                                                                                                                                                                                                                                                                                              364
                                                                                             9
                                                                              212 YGGLGSQGAGRGGLGGQGAG--AAAAAGGAGQGGLG----GQGAGAAAAAAGGAGQGGYG
                                                                                                                                                                                                                                                                       34 AAAAAGGAGGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGLGGQG
                                                                                                                                                                                       152 AAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGGQGAGQGG
                                                                                                                                                                                                                                                                                                                                  AAAAAAG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCDVLIQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----GQGAGAAAA------AAAAAAGGVRQGGYGGLGSQGAGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Aranese;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
NCBI_TaxID=6915;
                               114;
                                Indels
   Length
                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GAGQGGI,GGQGAGAAAAAGGAGQ----GGLGGQGA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9024AF2F65C5F75A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-UN-2002 (TrEMBLrel. 21, Last annotation update)
Dradline slik protein (Fragment).
Nephila clavipes (Orb Spider).
Score 2313; DB 5;
Pred. No. 2.1e-110;
1; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Submitted (NOV-2001) to the EMBL/GenBank/DDBJ
EMBL; AY061814; AAL32375.1; -.
INTECPPO; IPRO00104; ANTIFREEZEI.
PRINTS; PRO0308; ANTIFREEZEI.
NON_TER 644 644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               644 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               52104 MW;
85.7%;
ilarity 80.1%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                644 AA;
                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               602
              Best Local Sim:
Matches 482;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LL 521
```

```
Q9BIU3
Q9BIU3;
                                                                                                                                62
                                                                                                                                                                                                                                           142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
                                                     g
                                                                                                                            a
                                                                                                                                                                                                     g
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCOCC OCT BY SET TO SE
                       ð
                                                                                        ò
                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                      10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26;
                                                                                                                                                                                                                      61 AGAAAAA----AGAGQGGYGGLGSQGAGRGGLGGQGAG--AAAAAGGVGQGGL 108
                                                                                                                                                                                                                                                                                                                                                                      GAG--AAAAAGGAGQGGLG----GQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAAA 282
                                                                                                                                                                                                                                                                                                                                                                                                                                             AAAAGGAQQGGYGGQGGGYGGLGSQGAGRGGLGGQGAGAAAAA------329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGVGAGGLGGQGAGAA 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                        Gaps
                                                                      6 GGQQMGRIRIR-----GYGGLGGQGAGQGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQG 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Argiope trifasciata.
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Araneidae; Argiope.
NCBI_TaxID=156845;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Silk
                                      64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 58.2%; Score 1570.5; DB 5; Length 648; Best Local Similarity 60.3%; Pred. No. 1e-72; Matches 379; Conservative 26; Mismatches 91; Indels 133;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-21179804; PubMed=11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
"Extreme Diversity, Conservation, and Convergence of Spider: Fibroin Sequences.",
Science 291:260-2605(2001).
EMBL: AF350266; AAK30595.1; -.
 Length 644;
                                      16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 648 AA; 52887 MW; ACDB67067B2302CB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUN-2001 (TrEMBLrel. 17, Last sequence update) 01-DEC-2001 (TrEMBLrel. 19, Last annotation update) Major ampullate spidroin 1 (Fragment).
Score 2290; DB 5;
Pred. No. 3.2e-109;
4; Mismatches 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          648 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last seq
01-DEC-2001 (TrEMBLrel. 19, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VVSQIGASNPGLSGCDVLIQALL 521
84.8%;
85.1%;
                                      Matches 479; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9BIU7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NON TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MASP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9BIU7
                                                                                                                                                                                                                                                                                             169
                                                                                                                                                                                                                                                                                                                                239
                                                                                                                                                                                                                                                                                                                                                                    229
                                                                                                                                                                                                                                                                                                                                                                                                         568
                                                                                                                                                                                                                                                                                                                                                                                                                                             283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09BIU7
                                                                                                             g
                                                                                                                                               ò
                                                                                                                                                                                a
                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                    ογ
                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
192
                                                                                                                                                                                                                                                                                                                             239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 433
                                                                                                                     121
                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                421
                                                                                                                                                                                                                                                                                                                                                                                                                                                     361
  43
                                        61
                                                                                 44 G-GYGGLGSQGAGRGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAA--- 99
                                                                                                    RGGSGGGAGAAAAAAGGAGGGGYGGLGSQGAGRG---GLGGQ-----GAGAAAAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  378 AAGGVRQGGYGGLGSQGAGRGGQGAGAAAAA----GGAGQGGTGGLGGQGVGAGGLGGQ
                       100 -----AGGVGQG---GLGGQ-GAGQG--AGAAAAAGG-AGQGGYGGLGSQGAG
                                                                                                                                                                                                                                                                                   G-AGQGGYGGLGGQGAGQGGYG------GLGSQGAGRGGLGGQGAGA--AAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                               323 GAAAAAA - AGGAGQGGLGGQGAGAAAAAGGAGG----GLGGQ-GAGQGAGAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          434 GAGAAAAVGAGGGCYGGVGSGASAASAASRLSSPQASSRVSSAVSNLVAS-GPTNSAAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240 GQGGLGGQGAGAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAA--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --AGGAGQG---GYGGQG------AGQGGYGGLGSQGAGRGGLGGQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AGGOGGOGGYGGLGSOGAGOGGYGOGGAAAAAAAAAAGGOGGOGGYGGLGSOGAGOGGYGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Fibroin 2 (Fragment)
Dolomedes tenebrosus.
Eukaryota, Metazoa, Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Eltelegynae; Lycosoidea; Pisauridae; Dolomedes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE=21179804; PubMed=11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Extrame Diversity, Conservation, and Convergence of Spider
"Extrance Sequences.";
Science 291:2603-2605(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                56176 MW; CB46EF9F1991F33E CRC64;
ASMTGGQQMGRIRIRGYGGLGGQGAGQG----AGAAAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; AF350270; AAK30599.1; -.
InterPro; IPR000104; Antifreeze_1.
InterPro; IPR002952; Eggabell.
PRINTS; PR01298; ANTIFREEZEI.
PRINTS; PR01228; EGGSHELL.
NON_TER
SEQUENCE 691 AA; 56176 MW; CB466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
```

Gaps --G 48

Indels 212; Length 988;

143

195

530 286

------GQGAGAAAAA

650 361 398

```
234 AAGAGGYGRG---AGGYGGQGGYGAGAGAGAAAAAGAGAGGAGGYGRGAGAGAGAAGAG 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          711 YGAGAGAAAAAAGAGGAAGYSRGGRAGAAGAGAGAAAGAGAGAGAGGYGGOGGYGA--- 767
                                                                                                                                                                                                                                            OGAGOGAGAAAAAAAAAAG-------GVRQGGYGGLGSQGAGRG
                                                                                                                                                                          49 LGSQGAGRGGQ-----GAGAAAAAAGGAGQGGYGGLG------SQGAGRGGLGGQ--
                                                                                                                                                                                                                                                                                144 -----GSGGQGAGAAAAAGGAGGGYGGLGSQGAGRGGLGGQGAGA-AAAAAAGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                            GGAGQGGYGGQG------GLGSQGAGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 826 AAAGAGAGAGAGAGAGGAGGYGGQGGYGAGAGAAAAAGAGAGGYGDKEIACWSRCRYT
                                                                                                                                                                                                                                                                                                                                                                                                        GLGGQGAGAAAAAAAGGAGQGGLG-GQGAGAAAAAGGA-------GQGGLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----ASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCDVLI
                                                                                                                                                                                                                             ----GAGAAAAAGGVGQGGLGGQGAGAGAAAAAAGGAGQGGYGGLGSQGAGRG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GQGAGAAAAAAGGAGQGGTGGLG------GAGVGAGGLGGO-----GAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       196 QGGYGGLG-GQGAGQGGYGGLGSQGAGRGGLGGQGAGA-AAAAGGAGQGGLGGQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-JUN-2001 (TrEMBLrel. 19, Last annotation update)
Fibroin | Fragment).
Dolomedes tenebrosus.
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae; Araneomorphae; Entelegynae; Lycosoldea; Pisauridae; Dolomedes.
                                                                                                                    2 ASWTGGQQMGRIRIRGYGGLGGQGAGQGAGAAAAAAAGAGGGGGGGG----
                                         461E03DF53F7085D CRC64;
                                                                  Ouery Match 51.7%; Score 1397; DB 5;
Best Local Similarity 47.7%; Pred. No. 9.4e-64;
Matches 345; Conservative 23; Mismatches 144;
                                                                                                                                                                                                                                                                                                                                                                                       249 ----AGAAAAAGGAGQGGYGGLGSQGAGRG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           854 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                         988 AA; 79082 MW;
    InterPro; IPR000817; Prion.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
             PRINTS; PR00341; PRION.
NON_TER 1 1 1 SEQUENCE 988 AA; 790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QVLL 949
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              518 QALL 521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9BIU4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9BIU4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             462
                                                                                                                                                                                                                              93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         946
                                                                                                                                                                                                                                                                                                                                                                                                                                            287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 7
Q9BIU4
    STRE
                                                                                                                                                g
                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                       Op
                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                              å
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                           34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [1]
SEQUENCE FROM N.A.
MEDLINE-98200471; PubMed-9541398;
COLGIN M.A., Lewis R.V.;
"Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like 'spacer regions'.";
Protein Sci. 7:667-672(1998).
EMBL; AF027735; AAC14589.1; -.
                                                                                                                                                                          456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 636
                                                                                                                    247
                                                                                                                                                                                                                                                                               298
                                                                                                                                                                                                                                                                                                                                    291 GQGAGAGAAAAAAGGAGAGAGGGYGQGGYGQG-GGAGAAAAAAAAGGGSGSGQGGYGGQG 349
                                                                                                                                                                                                                                                                                                                                                                           343
                                                                                                                                                                                                                                                                                                                                                                                    350 -GLGGYG----QGA-----GAGAGAAAAAAGAGSGQGGYGGQGGLGGYGQGAGAAA 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         516
                                                                                                       65 AAAAG--GAGQGGYGGLGSQGAGRGGLGGQGAG-----AAAAAAGGV-----GQG 106
                                                                                                                                                          GLGGQGAGQAGAAAAA---GGAGQGGYGGLGSQGAGRGGSGGQGAGAAAAAGGA--G 161
                                                                                                                                                                                                                            391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AASAAAAGGAGAGAGGYGGGCGCGGYGQGAGAGASAAAASASGAGSGQGGYGGQGGYGGGGT 576
                            Gaps
                                                               8
                                                     -SQGAGRG-----GQGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GAGAASSAGVAVTVGNTVSRLSSPQAASRVSSAVSSLVSNGQVNVAALPSIISSLSSSIS
                                                                                                                                                                                                                                                                 201 GLGGQGAGQGGYGGLGSQGAGRGGLG-GQGAGAAAA-----GGAGQGGLGGQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                         AA-----VGAGQGGYG-------GVGSGASAASAAA-----VGAGQGGYG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------SRLSSPQASSRVSSAVSNLVASGPINSAALSSTISNVVSQIG
                                                                                                                                                                                                              QGGYGGLGSQGAGRGGLG-GQGAGAAAAAAGGA-----GQGGYG------
                                                                                                                                                                                                                                                                                                                    ----GAGAAAAAAG---GAGQGGYGGLGSQGAGRGGQGAGAAAAAA----GGAGQGGYGGQG
                                                                                                                                                                                                                                                                                                                                                                        AGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAG-GAGQGGLGGQG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SQGAGRG----GQGAGAAAAAAGGAGGGTGGTG-----GQGVGAGGLG--GQGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                            AGAAAAAAGGAGGGGGGGGAGAAAAAAAAAAAAGGVRQGGYGGLG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Araneae;
Nephila.
                           203;
Length 691;
                           82; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae;
NCBI_TaxID-6915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               017434 PRELIMINARY; PRT; 988 AA.
017434;
01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
Minor ampullate slik protein MiSp1 (Fragment).
Nephila clavipes (Orb spider)
 DB 5;
               , ee-64;
; Score 1405.5;

Pred. No. 2.6e
29; Mismatches
                                                  GQGAGQGAGAAAAAAGGAGQGGYGGLG-----
52.1%;
53.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         521
                           363; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ||: | |:||:| || ASSTAASDCEVLVQVLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ASNPGLSGCDVLIQALL
            Similarity
           Best Loca
Matches
                                                  23
                                                                                                                               61
                                                                                                                                                          107
                                                                                                                                                                                                              162
                                                                                                                                                                                                                                       179
                                                                                                                                                                                                                                                                                                                    248
                                                                                                                                                                                                                                                                                                                                                                        299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        505
                                                                                                                                                                                                                                                                                                                                                                                                                            344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     017434
                                                                           g
                                                                                                       ò
                                                                                                                              셤
                                                                                                                                                          à
                                                                                                                                                                                  셤
                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
```

S

```
49;
                                                                                                                                                                                                                                                                                                             123 SATAPAAGGSGFGQGGFGNRGGKGAYGQSAGAGVGAAATAAAGGAGSGQGGYGDQGLGG 182
                                                                                                                                                                                                                                                                                                                                                                                                               243 QGGLGGYGQGAGAGAAAAAGGGGGYGQGYTGDQG-GRGGYGQGSGAGSATAAAAGGAGF 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                    202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  264 YGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGQGAGQGGY---GGLGS--QGAG---- 313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        539 AASAGGAGSGRGGYGDQGGLGGYGQGAGAASAAAGGGDGYGQGGYGURGYGSYGQGA 598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           836
                                                                                                                                                                                                                                                         105 QGGLGGQGAGQAGAAAAAAGGAGQGGYGGLGSQGAGRGGSG-GQGAGAA-AAAAGGA-- 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LG--GQGAG----AAAAAGGAG-----QGGLG----GQGAGAAAAAAAGGA---GQGG 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -------RGGLG--GQGAGAAAAAAGGA---GQGGLG------GQGA 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              393 QGA-GRG-GQGAGAA-AAAAGGA------GQGGTGGL--------GGQGVGA 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                428 GGLG------GQGAG--AAAA-----VGAGQGGYG-----GVGSGASAASAA--- 461
                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                           6 GGQQMGRIRIRGYGGLGGQGAGQAAAAAAAAAGGAG------QGGYGGLGSQGAG-- 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       599 GAGAAATSAAGGAGSGRGGYGEQGGLGGY--GQGAGAGAAASTAAGGGDGYGQGGYGNQGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                183 YGQGAGAGAASAAAGGGDGYEQGGYGNQGGLGSFGQGAGAAAAAAAAAAAAGGAGSGRGGYGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                   161 GQGGY--GGLGSQGAGRGGLG-GQGAGAAAAAAGGAG------QGGYGGL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----ASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCDV
                     SEQUENCE FROM N.A.
MIDLINE-21179804; PubMed-11283372;
MATCHINE-21179804; PubMed-11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
Extreme Diversity, Conservation, and Convergence of Spider Silk
Fibroin Sequences ";
Science 291:2603-2605(2001).
EMBL; AF350269; AAK30598.1; -,
Interpro: IPR000217; Tubulin.
PROSITE: PS000277; TUBULIN; UNKNOWN_1.
                                                                                                                                                                                         Query Match
48.4%; Score 1305.5; DB 5; Length 854;
Best Local Similarity 48.6%; Pred. No. 3.6e-59;
Matches 382; Conservative 29; Mismatches 98; Indels 277;
                                                                                                                                                                 5C52F1B0AE31C6C2 CRC64;
                                                                                                                                                                 71881 MW;
                                                                                                                                                      854
                                                                                                                                                                 854 AA;
NCBI_TaxID=156846;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    516 LIQALL 521
                                                                                                                                         NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            345
                                                                                                                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 462
 OX
RRY
RRT
RRT
DR
DR
FFT
SQ
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
```

```
20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GQGGYGGLGSQGAGRGGLG-GQG----AGAAAAAGVGQGGLGGQGAGQGAGAAAAAGG 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             186 AAAAAAGGAGGGGGGGGGGGGGGGGT----GGLGSQ-----GAGRGGLGGQGAGAAA 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              235 AAGGAGQGGLGGQGAGAAAAAGG-AGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGG 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              294 YGGQGAGQGGYGGLGSQGAGRGGLG-GQ-GAGAAAAAAAGGAGQGGLGGQGAGAAAAAA 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGAGQG----GLGGQ-GAGQGAGAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAA 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          408 AAGGAGQGGTGGLGGQGVGAGGLGGQGAGAAAVGAGQGGYGGVGSGASAASAASRLSS 467
                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 GGQGAGQGAGAAAAAAAGGAGQGGYG-----GLGSQGAGRGGQGAGAAAAAAGGA 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 GQGGAGAAAAAAAAAAAGGAGGAGGAGQGYGSGLGGGGGAGAAAAAAAAAGGQG 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Argiope aurantia.
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoldea; Araneidae; Argiope.
                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 108;
                                                                                                                                                                                                                                                                    Spider
                                                                                                                                                                                                                                                                                                                                                                                                                     Length 447;
                                                                                                                                                                                                                                SEQUENCE FROM N.A.

MEDLINE-21179804: PubMed-11283372;

MEDLINE-21179804: PubMed-11283372;

Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;

"Extreme Diversity, Conservation, and Convergence of Spic Fibroin Sequences.";

Socience 291:2603-2605(2001).

SEMBL; AF350262: AAK30591.1; -.

InterPro: IPR000817; Prion.

PRINTS: PR00341; PRION.

PRINTS: PR00341; PRION.

NOW_TER 47 47

SEQUENCE 447 AA; 35630 MW; 78D160B7E6DF24C7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                       447 AA; 35630 MW; 78D160B7E6DF24C7 CRC64;
                                                                                            01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
Major ampullate spidroin 1 (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                             85;
                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
46.5%; Score 1256; DB 5;
Best Local Similarity 58.6%; Pred. No. 6.7e-57;
Matches 303; Conservative 21; Mismatches 85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PQASSRVSSAVSNLVAS-GPTNSAALSSTISNVVSQI 503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     447 AA
                                                                     PRT;
                                                                   PRELIMINARY;
                                                                                                                                                                                                        NCBI_TaxID-156844;
837 LVQVLL 842
                                                                                Q9BIV1
                                                                  Q9B1V1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            389
                                         RESULT 8
                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            입
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

RESULT 9

```
Nephila madagascariensis.
                                                                                                                                                                                                                                  Best Local Similarity 49.9
Matches 284; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                     563 AA;
                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                  Fibroin Sequences.
Science 291:2603-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUN-2001
01-JUN-2001
01-JUN-2001
                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           O9BIT7
                                                                                                                                                                                                                                                                                                           17
                                                                                                                                                                                                                                                                                                                               65
                                                                                                                                                                                                                                                                                                                                                                                             186
                                                                                                                                                                                                                                                                                                                                                                                                                                      246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
  DAPPP
PPPPP
                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                           å
                                                                                                                                                                                                                                                                                                                              D
                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                            26;
                                                                                                                                                                                                                                                                                                                                       AAAAAAGGGGGGGGGGGGGGGGGGGGG---AAAAAGGVGQG-----GLGG 110
                                                                                                                                                                                                                                                                                                                                                                                156 QGPYGPGASAAAAAGGYGPGSGQGPGQQGPGGQGPYGPGASAAAAAGGYGPGSGQQGP 215
                                                                                                                                                                                                                                                                                                                                                                                                                                       315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      330 QGPGQQGPGGQGPYGPGASAAAAAAGGYGPGSQQGPGQQGPGQQGPGQQGPGQQGPGQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QGAGAAAAAAAAAAGGVRQG-GYGGLGSQGAGRG---GQGAGAAAAAAGGAGQG-GTGGL 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  494
                                                                                                                       SEQUENCE FROM N.A.

MEDLINE-96178678; Pubmed-8600519;

MEDLINE-96178678; Pubmed-8600519;

Guerette P.A., Glazinger D.G., Weber B.H., Gosline J.M.;

Sallk properties determined by gland-specific expression of a spider fibroin gene family.";

Science 272:112-115(1996).

EMBL, U47855; AAG47010.1;

InterPro; IPR001419; Glutenin.

PRINTS; PR00210; GLUTENIN.
                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                         7 GOQMGRIRIRGYGGLGGQGA-GQGAGAAAAAAG---GAGQGGYGGLGSQGAGRGGQGAG 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100 AAAAAAGGNGPGS----GQQGAGQQGPGQQGPGASAAAAAAGGYGPGSGQQGPGQ
                                                                                                                                                                                                                                                                                                                                                                                                                          GSQGAGRGCLGGQG----AGAAAAAAAGGAGQG-----GYGGLGGQG-----AG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  209 QGGYG-GLGSQGAGRGGLGGQG-----AGAAAAGGAGQGGLGGQGAGAAAAAGGAGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           262 GGYGGLGSQGAGRGGQGAGAAAAAAG----GAGQGGYGGQGAGQGGYG--GLGSQGAGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLGGQGAGAAAAAAAGGAGQGGLGGQ----GAGAAAAAAG----GAGQGGLGGQGAG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 443 QGPG------QQGPGQQGPGQQGPGQQGPGQQGPYGPGASAAAAAGGYGPGSGQQGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGQGVGAGGLGGQG--AGAAAAVGAGQGGYGGVGSGASAASAAASRLSSPQASSRVSSAV
                                                                                     Araneae;
                                                                                                                                                                                                                                                                           98;
                                                                                                                                                                                                                                                       Length 636;
                                                                      Araneus diadematus (Spider).
Eukaryota: Metazoa; Arthropoda; Chelicerata; Arachnida; Ara
Araneomorphae; Entelegynae; Araneoidea; Araneidae; Araneus.
                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                 636 AA; 56074 MW; E2F41180C978BB59 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                     Query Match 40.7%; Score 1099; DB 5; 18est Local Similarity 52.7%; Pred. No. 8.4e-49; Matches 307; Conservative 20; Mismatches 158;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCDVLIQALL
636 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        563
                     Created)
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT';
                    01,
01,
                             01-NOV-1996 (TrEMBLrel. 01, 01-JUN-2001 (TrEMBLrel. 17,
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                   (TrEMBLrel.
                                                   Fibroin-3 (Fragment).
                                                                                                       NCBI_TaxID=45920;
                    01-NOV-1996
                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9BIT5
Q9BIT5;
                                                              ADF-3
                                                                                                                                                                                                                                                                                                                                                                                                                        169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        390
                                                                                                                                                                                                                                                                                                                                        63
                                                                                                                                                                                                                                                                                                                                                                                 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9BIT5
ò
                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGLGSQGAGRGGLGGQGAGAAAAG-GVGQGGLGGQGAGAGAAAAAAGGAGQGGYG-- 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29 GAGAAAAAAAGGAGQ---GGYG-----GLGSQGAGRGGQGAGAAAAAGGAGQGGY 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 || | | | | | | | PGQQGPGGYGPGQQGPSGPGSAAAAA
PGQQGPGGYGPGQQRPSGYGPGQQGPSGPGSAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  134 --GLGSQGAGRGGSGGQGAGAAAAAAGGA--GQGGYGGLGSQGAGRGGLGGQG----AGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180 AAAAAAAGPGQGFGGYGPGQQGPGGY-----GPGQQGPSGPGSAAAAAAAAAGPGQQGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAAAAAAAGGAGQGGYGGQGAGQGGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----QGPGQQGPSGPGSAAAAAAAAGPGPGGYGPGQYGPGGYGPSGPGSAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             354 AGGGGLGGQGAGQ------GAGAAAAAAAAAAGGVRQGGYG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLGSQG-AGRGCQGAGAAAAAGGAGQGGTG----GLGGQG--VGAGGLG-----G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QGAGAAAAVGAGQGGYGGVGSGASAASAASRLSSPQASSRVSSAVSNLVASGPTNSAAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLGSQGAGRGGLGGQGAGAAAAAAAGGAGQGGLG------GQGAGAAAAAAGG
                                                                                                                                                                                                                                                               MEDIINE-21179804; PubMed-11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
"Extreme Diversity, Conservation, and Convergence of Spider Silk
Fibroin Sequences."
                                                                                                                                             Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
NCBI_TaxID=115969;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38.6%; Score 1042.5; DB 5; Length 563; 49.9%; Pred. No. 5.6e-46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                         48361 MW; OAEAD885304B8B6B CRC64;
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
Major ampullate spidroin 2 (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (TrEMBLrel. 17, Created)
(TrEMBLrel. 17, Last sequence update)
(TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21; Mismatches 155;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SSTISNVVSQIGASNPGLSGCDVLIQALL
                                                                                                                                                                                                                                                                                                                                            colence 291:2603-2605(2001).
EMBL: AF350278; AAK30607.1: -.
```

```
MEDLINE-21179804; PubMed-11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
"Extreme Diversity, Conservation, and Convergence of Spider
Fibroin Sequences.";
                                                                                                                                                                                                                                                                                                Length 360;
                                                                                                                                                                                                                                                                        360 AA; 30253 MW; 0EB6E46EC1122740 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Fibroin.
                                                                                                                                                                                                                                                                                                                           48;
                                                                                                                                                                                                                                                                                                 DB 5;
                                                                                                                                                                                                                                                                                             Query Match 37.1%; Score 1003; DB 5; Best Local Similarity 54.5%; Pred. No. 3.8e-44; Matches 247; Conservative 24; Mismatches 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              489 SAALSSTISNVVSQIGASNPGLSGCDVLIQALL 521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Antheraea pernyi (Chinese oak silk moth).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ampullate spidroin 1 (Fragment)
                                                                                                                                      Fibroin Sequences.";
Science 291:2603-2605(2001).
EMBL, AF350273; AAX30602.1;
InterPro; IPR000104; Antifreeze_1.
InterPro; IPR000817; Prion.
PRINTS; PR000817; Prion.
PRINTS; PR01228; EGGSHELL.
PRINTS; PR01228; EGGSHELL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                             Latrodectus geometricus.
                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     076786 076786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
  à
                                                                                                                                                                                                                                                                                                                                                                           硆
                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31;
                                                                                                                                                                                                                                                                                            1052 AAAAAGRGPG-----GYGPGQQGPGGPGAAAAAGPGGYGPGQQGPGAAAAAAAGRGPG 1106
                                                                                                                                                                                                                                                                                                                                     1512 GPGAAAAAAGTGPGGYGPGQQGPGGSGAAAAAGRGPGGYGPGQQGPGGPGAAAAAGP 1571
                                                                                                                                                                                                                                                                                                                                                                                                                           128 GQGGYG------GLGSQGAGRGGSGGQGAGAAAAAGGAGQGGYGGLGSQGAGRG 176
                                                                                                                                                                                                                                                                                                                                                                        ----GQGAGAAAAAG-----GVGQGGLGGQGAG-------QGAGAAAAAGGA 127
                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                       44
                                                                                                                                                                                                                                                                                                                       45 GYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYG----GLGSQG------AGRGGLG 90
                                                                                                                                                                                                                                                                     2 ASMTGGQQMGRIRIRGYG------GLGGQGAG-QCAGAAAAAAGGAGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           177 GLGGGGAGAAAAAAAGGAGGGYG----GLGGQG-----AGQGGYG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               249 AGAAAAAAGGAGGGYGGLGSQGAGRGGQGAGAAAAAAG----GAGQGGYGGQG----A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  300 GQGGYGGLGSQG-----AGRG----GLGGQGAGAAAAAAAGGAGQGGLG-GQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   343 GAGAAAAAAGGAGGGLG-GQGAGQGAGAAAAAAAAAAGGVRQG--GYGGLGSQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -GAGRGGQGAGAAAAAAGGAGQGGTG-------GLGGQGVGAGGLGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              214 --GLGSQGAGRGGLGGQGAAAAAG-----GAGOGGLGG-------OG
                                                                        SEQUENCE FROM N.A.

MEDLINE-21179804; PubMed=11283372;
Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;

"Extreme Diversity, Conservation, and Convergence of Spider Silk
"Extreme Sequences.";
Science 291:2603-2605(2001).

EMBL, AF350276; AAK30605.1;
                      Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Tetragnathidae; Nephila.
NCBL_TaxID=115969;
                                                                                                                                                                                                                   38.0%; Score 1025.5; DB 5; Length 1953; 46.8%; Pred. No. 1.2e-44; Live 12; Mismatches 135; Indels 189;
                                                                                                                                                                                          159383 MW; C82B4DACD043C3BD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
Major ampullate spidroin 2-like protein (Fragment)
Nephila madagascariensis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     435 AGAAAAVGAGQGGY -- GGVGSGASAASAAAS 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (TrEMBLrel. 17, (TrEMBLrel. 17, (TrEMBLrel. 18,
                                                                                                                                                                                                                                           Matches 295; Conservative
                                                                                                                                                                              1953
                                                                                                                                                                             1953 19:
1953 AA;
                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUN-2001
01-JUN-2001
01-OCT-2001
                                                                                                                                                                             NON_TER
SEQUENCE
                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                Local
                                                                                                                                                                 NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9BIU0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9BIU0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1632
                                                                                                                                                                                                                                                                                                                                                                        91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
g
                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95444
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
```

```
19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGYGGLGSQGAGRGGSGGQGAG-AAAAAAGGAGQGGYGGLGSQGAGRGGLGCQGAGAAAA 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         189 AAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRGCLGGQGAGAAAAGGAGQGGLGGQG 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -QGYGQGGAGQGGAAAAAAAAAGGAGGGYG-------198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              369 GAAAAAAAAAAAGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGTGGLGGQGVGAG 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---SAAAAAAAGGAGGGY---GRGGAGGGAGSAAAAAAAAAGGSGQGGQGGY-CQGYGQG 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        429 GLGGQGAGAAAAVGAGQGGYGGVGSGASAASAASRLSSPQASSRVSSAVSNLVASGPTN 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            70 GAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAGGVGQGGLGGQGAGGAGAAAAAAAGGAGQ 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86 AAAAGSGQG-----GGGGYG----QGYGQ-------GGAGQ---GGAG 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   249 AGAAAAAAGGAGGGYGGLGSQGAGRGGQGAGAAAAAGGAGGGGYGGQGAGQGGYGGLG 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38 GGOGGYG-QGYGQGGAGQGGAGAAAAAAGGAGQGGY------GRGGAGQGAAAA 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pteryyota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Saturniidae; Saturniinae; Saturniini; Antheraea.
NCBI_TaxiD-7119;
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
Araneomorphae; Entelegynae; Araneoidea; Theridlidae; Latrodectus
NCBI_TaxID-156851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 134;
```

```
54
                                                                                                                                                                                                                                                                          557
                                                                                                                                                                                                                                                                                                                                                                     190
                                                                                                                                                                                                                                                                                                                                                                                       736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         975
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1155
                                                                                                                                                                                                                                                                                             97
                                                                                                                                                                                                                                                                                                                                                                                                         231
                                                                                                                                                                                                                                                                                                                                                                                                                                              272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   300
   ó
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                               24;
                                                                                                                                                                                                                                               194
                                                                                                                                                                                                                                                                 400
                                                                                                                                                                                                                                                                                                      458
                                                                                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                                                                 636
                                                                                                                                                                                                                                                                                                                                                                                                                                    420
                                                                                                                                                                                                                                                                                                                                                                                                                                                     691
                                                                                                                                                                                                                                                                                                                                                           MEDLINE-20496953; PubMed-11040284;
Sezutsu H., Yukuhiro K.;
Sezutsu H., Yukuhiro K.;
Dynamic rearrangement within the antheraea pernyi silk fibroin gene
"Dynamic rearrangement within the antheraea pernyi silk fibroin gene
"Ja associated with four types of repetitive units.";
J. Mol. Evol. 51:329 338 (2000).
EMBL; AF083334; AAC32606.1;
SEQUENCE 2639 AA; 216057 MW; 2EE3310DEEB09B9A CRC64;
                                                                                                               Gaps
                                                                                                                                20 GLGGQGAGQG-----AGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQ 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PE_PGRS_family protein.
MT3612.1.
Wycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
[1]
                                                                                                                                          GGYGGLGSQGAGRGGLGGQGAGAAAA---AGGVGQG---GLGGQGAGQGAGAAAAA---
                                                                                                                                                                                 281 AGSGAGGSGGYGGYGSDSAAAAAAAAAAAAAGSSAGGAGGYGWGDGGYGSDSAAAAAAA
                                                                                                                                                                                                                                                            341 AAAAASSGAGGRGDGGYGSGGSSAAAAAAAAAAAAAARRAGHDRAAGSAAAAAAAAAA
                                                                                                                                                                                                                                                                                            ----GAGRG---GSGGQGAGAAAA---
                                                                                                                                                                                                                                             -----AGGAGGGYGGLGSQGAG------RGG----LGGQGAGAAAAAAGGA
                                                                                                                                                                                                                                                                                                                                                                                              AAAAAAAGGAGGGLGG-------QGAGAAAAAAGGAGQGGL------GG
                                                                                                                                                                                                                                                                                                                                                                                                                                  QGAGQGAGAAAAAAAAAAGGVR-QGGYGGLGSQGAGRGGQGAGAAAAAAAAGGAGQGGTGGL
                                                                                                                                                                                                                                                                                 GQGGYGGLGGQGAGQGGYG------GLGSQGAGRG-GLGGQG-----
                                                                                                                                                                                                                                                                                                                      230 ----AGAAAAAGGAGQGGLGGQG-----AGAAAAAAGGAGQGGYGGLGSQGAGRGGQG
                                                                                                                                                                                                                                                                                                                                                                                                                AAAAAAAAASGAGGSGGGYGWGDGGYGSDSAAAAAAAAAAAAAAGSGAGGGYGWGDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGQGVGAGGLGGQGAGAAAAV-----GAGQGGYGGVGSGASAASAASRLSSPQASSR
                                                                                                              Indels 164;
                                                                                            Length 2639;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                      752 HDSAAGSAAAAAAAAAAAAAS-----GAGGSG 778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           474 VSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPG 509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 1715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                        ---GGAGQGGYGGLGSQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20,
20,
                                                                                                    Best Local Similarity 42.6 Matches 271; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q8VIZO;
01-MAR-2002 (TrEMBLrel.
01-MAR-2002 (TrEMBLrel.
01-JUN-2002 (TrEMBLrel.
                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OSINSO
                                                                                                                                                                                                                                                                                                                                        459
                                                                                                                                                                                                         125
                                                                                                                                                                                                                                                                                                                                                          279
                                                                                                                                                                    74
                                                                                                                                                                                                                                             157
                                                                                                                                                                                                                                                                                 195
                                                                                                                                                                                                                                                                                                                                                                           517
                                                                                                                                                                                                                                                                                                                                                                                              325
                                                                                                                                                                                                                                                                                                                                                                                                                577
                                                                                                                                                                                                                                                                                                    401
                                                                                                                                                                                                                                                                                                                                                                                                                                   362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q8VIZ0
셤
                                                                                                                                                                                    g
                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                    à
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
```

```
1035 VGGDGGEGASGLGLGLSGFDGGQGGQGGAGGNAGAGGINGAGGAGGTGGAGGDGAPATLI 1094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GQGGYGGLGSQG-----AGRGGLGGQ------GAGAAAAA 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             53
SEQUENCE FROM N.A.
STRANGCO [1551 / OSHKOSH;
Fleischmann R.D. Alland D., Eisen J.A., Carpenter L., White O.,
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Fleterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.I
Belcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Blishal W.;
"Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 ASMTGGO--QMGRIRIRGYGGLGGQGAGAGAA----AAAAAGGAG-QGGYGGLGSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A-----GRGGGGAGAAAAAGGAGA------GGYGGLGSQGAGRGGLGGQGAGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GNTGVGGTNGSGGQGGTGGAGGAGGAGGVGADNPTGIGGTGGTGGKG-GAGGAGADA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       616 TATGATGGTGFAGGAGGAGGAGGGSSGAGGTNGSGGAGGAGGAGGAGGAGADNPTGIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 141 GRGGSGGQGAGAAAAAGGA-GQGGYGG-LGSQG-----AGRGGLGGQ-GAGAAAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        855 FDGGGGGGGGGGGGAGGINGAGGNGGDGGTGAAGLGDNGGVGGDGGAGGAAGNGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  915 NAGVGLTAKAGDGGAAGNGGNGGAGGAGGAGDNNFNGGQGGAGGQGGQGGLGGASTTSIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1095 GGPDGGDGGQGGIGGDGGNAGFGAGVPGDGGDGGNAGFGAGVPGDGGIGGTGGAGGAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---AAGGAG-QGGYGGLGGQGA------GQGGYGGLGSQGAGRGGLGGQGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --AGRGGQGAGAAAAAGG----AGQGGYGGQGA-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----GAGGLGGQGAGAAAVGA--------GQGGYGGVGSGASAASA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              387 YGGLGSQGA-----GRGGQGAGAAAAAGG----AGQGGTGGLGGQGV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GADGDPSIDGGQGGAGGHGGQGGKGGLNSTGLASAASGDGGNGGAGGAGGNGGD 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA - - AAGGV - - - - GQGGLGGQGAGQGAG - - - AAAAAAGGAGQGGYGGLGSQGA - - - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1715;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  203;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AASRLSSPQASSRVSSAVSNLVASGPTNSAALSSTISNVVSQIGASNPGLSGCD 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ProDom; PD001223; PE_region; 1.
PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_4.
SEQUENCE 1715 AA; 133832 MW; E857B2A774C771B0 CRC64;
                                                                                                                                                                                                                                            laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AE007163; AAK47971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --GAAAAAGGAGQGGLGGQ----GAGAAAAAAGGAGQGGYGGLGSQG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 34.9%; Score 942.5; DB 16; Best Local Similarity 37.4%; Pred. No. 1.7e-40; Matches 267; Conservative 20; Mismatches 224;
                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000084; PE_region.
InterPro; IPR002173; PfkB.
Pfam; PF00934; PE; 1.
```

```
23;
                                                                                                                                                                                                                                                                                                                                                                                                            1479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1600 AAAAAAAAAAAASGAGGSGGGYGWGDGGYGSDSAAAAAAAAAAAAAGGAGGAGGYGGYG 1659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               206 ------AGGAGGYGGLGSQGAGRGGLGGQGAGAAA-----AGGAGQG 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           278 G-----AGAAAAAAGGAGGGYGGQG------AGQGGYGGLGSQGAG 313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           67 AAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAGGVGQG---GLGGQGAGQGAGAAAAA 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               152 -AAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAA----GGAGQGGYGGLGGQ 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------GQGAGAAAAGGAGGGGTGCLGSQGAGRGGQ 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      445 QGGYGGVGSGASAASAASRLSSPQASSRVSSAVSNLVASGPTNŞAALSSTISNVVSQIG 504
                                                                                                                                                                                                                                                                                                                          Query Match 34.9%; Score 941; DB 5; Length 2655; Best Local Similarity 40.6%; Pred. No. 3e-40; Matches 271; Conservative 24; Mismatches 179; Indels 194; Gaps
                                                                                                    Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterryota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysla; Bombycoidea; Saturnildae; Saturniinae; Saturniini; Antheraea.
                                                                                                                                                                         SEQUENCE FROM N.A.
TRANSPOSON-MARINER LIKE ELEMENT;
Lee J.-S., Hwang J.-S., Kang S.-W., Suh D.-S., Jin B.-R., Kim Y.-S.,
Lee K.-S., Goo T.-W., Yun E.-Y., Kim K.Y.;
"Complete nucleotide sequence of fibroin gene cloned from Antheraea
                                                                                                                                                                                                                                                                                                                                                                                    17 GYG-GLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQG-----AGAAAA 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     124 -----AGGAGQGGYG------GLGSQGAGRG-GSGGQGAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1420 GYGSGSSAAAAAAAAAAAAARRAGYDRAAGSAAAAAAAAAAAAAAAAGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                  yamama1.";
Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AF325500; AAK83145.1; -.
SEQUENCE 2655 AA; 218672 MW; 89BB810EF9CC4CA6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  314 RGGLGGQ---GAGAAAAAAGGAGQGGLGGQGAG------
                                             Last sequence update)
Last annotation update)
   2655 AA
                                Created)
   PRT;
                           01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19,
   PRELIMINARY;
                                                                                        Antheraea yamamai.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ASNPGLSG 512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 243 GLG-----
Q964F4
Q964F4;
                                                                         Fibroin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         505
                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
```

Db 1762 ASGAGGSG 1769

Search completed: December 18, 2002, 16:20:59 Job time : 36.9992 secs

This Page Blank (uspto)

~

Title: Perfect score:

Run on:

Scoring table: sednence:

Searched:

Database

```
Human protein sequence Novel human diagno Novel human diagno Novel human diagno Novel human diagno Aeromonas hydrophi Aeromonas hydrophi Omph protein of Pa Class II outer mem Mature class 2 por Neisseria meningit Cmph protein of Pa Class II outer mem Mature class 2 por Neisseria meningit Cmph protein of Pa Neisseria meningit Cmph protein of Pa Neisseria meningit Meningococcal groundingseria meningit Meningseria meningit Meningseria meningit Meningseria meningit Meningseria meningit
Novel human diagno
Novel human diagno
Novel human diagno
Novel human diagno
C. bifermentans DP
Sequence of ompK36
                                                                                                                        Human secreted pro
Novel human diagno
Novel human diagno
Novel human diagno
                                                                                 E. coll outer memb
E. coll outer memb
Novel human diagno
                                                                                                                                                                                                                                                                                                                                                                                                                                         P3 gene product of
Gonococcal porin-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  influenzae prot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Recovering structural polypeptides in a biological sample, useful for purifying and spinning spider silks and other structural proteins,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Outer membrane protein; ompF; structural protein; purification.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                                                                                                                                                                        AAE09799
AAY26025
AAY26023
AAR06038
AAR70763
AAW21742
                                                                                                                                                                                                                                                                                                                                                                   AAW21743
AAY26024
AAW21744
AAR70762
                                                                                                                                                                                                                                                                                                                                                                                                                        AAW21741
AAR07043
AAR69607
AAR66879
                                                                                                                                                                            AAM25524
ABG29003
                                                                                                                                                  ABG25865
                                                                                                                                                                                                                                                                                                                                                          AAR70764
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          E. coli outer membrane protein ompF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Butler
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB82612 standard; Protein; 341
 01-NOV-2000; 2000WO-US30086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20-JAN-2000; 2000US-0490291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŝ
                                                                                                                                                                          dello CM, Arcidiacono
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (MELL/) MELLO C M.
(ARCI/) ARCIDIACONO S.
(BUTL/) BUTLER M M.
(USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VPI; 2001-483136/52,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Escherichia coll.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200153333-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26-JUL-2001.
152.5
150.5
147.5
144.5
140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB82612;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB82612
ID AAB8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 1
 E. coll outer memb
E. coll wild-type
E. coll ompF porin
E. coll outer memb
E. coll outer memb
E. coll outer memb
E. coll growth and
Escherichia coll o
E. coll outer memb
E. coll outer memb
                                                                                                        (without alignments)
1995.696 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                     2. (SIDS2/gcgdata/geneseqp-embl/AA1980.DAT:*
2. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
4. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
5. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
5. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
6. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
7. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
8. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
9. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
11. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
12. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
14. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
15. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
16. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
17. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
18. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
18. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
19. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
19. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
19. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
19. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
20. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
21. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
22. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
22. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
23. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
24. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
25. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
27. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
28. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
28. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
29. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
20. (SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                                                                                          December 18, 2002, 16:05:10 ; Search time 22.7682 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                             US-09-490-291-9
1809
1 AEIYNKDGNKVDLYGKAVGL.....NKLGVGSDDTVAVGIVYQFA
            GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                  Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                           908470 seqs, 133250620 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AABB2612
AAY42550
AAW92998
AAY34058
AAY34058
AAG98980
AAG98980
AAG98980
AAG98980
AAG98980
                                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                    A_Geneseq_101002:*
                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                341
362
340
367
367
367
377
377
377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0
99.8
99.1
58.1
58.1
58.1
57.2
57.2
57.1
56.9
```

1809 1805 1792 1051 1051 1031 1033 1029

1264507861

Result No.

```
ö
comprises treating the sample containing the polypeptides with an acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 WEXNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   membrane proteins. The invention provides methods for purifying structural proteins. Organic acids are used to lyse recombinant cells or other biological samples (such as non-recombinantly derived cells), and enrich the purity and yields of structural proteins by hydrolysing many of the macromolecules while leaving proteins by hydrolysing many of the macromolecules while leaving perites of E. coll were lysed in 9.2 N valeric acid for 1 hr at room temperature. The cell lysate was clarified by centrifugation and applied to an SDS polyacrylamide gel for electrophoresis. The OmpF protein was then blotted onto a nitrocellulose membrane for N terminal sequencing. The simple extraction procedure yielded for the protein was then blotted onto a nitrocellulose membrane for N terminal sequencing. The simple extraction procedure yielded for the protein man approximately 85% purity. The new method has the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacteria; attenuation; deletion; mutant; vaccine; immune response; Gram negative; infection; diarrhoea; food poisoning; typhoid; salmonellosis; gonorrhoea; gastroenteritis; whooping cough.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     following advantages over prior art: it involves fewer steps, requires less time and smaller volumes of reagents, results in setter recovery of protein at higher purity (70-99%), is easy to scale up, and fibres can be spun in an environmentally benign solution reducing hazardous waste accumulation and cost.
                                                                                         present sequence is that of the Escherichia coli ompF outer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301 YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQFA 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301 YYENKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQFA 341
                                                    Claim 2; Page 43-44; 49pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY42550 standard; Protein; 362 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     E. coli wild-type ompF protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-DEC-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 341; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                     341 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     W09949026-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY42550;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 2
AAY42550
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
```

```
C (AAZ22889-Z22893) to produce a non-reverting deletion mutation. The mutant compt gene (AAZ22889) was used in the production of a bactcrium cartenuated by a non-reverting mutation in each of the compt gene, the arcc gene (AAZ22885) AAZ22884), and the compt gene (AAZ22885, AAZ22886). The arcc gene (AAZ22885) AAZ22884), and the compt gene (AAZ22885) AAZ22886). The arcc gene (AAZ22885) AAZ22886), and the compt gene (AAZ22885). The arcc c gene (AAZ22885) AAZ22886). The arcc c gene (AAZ22885) AAZ22886). The arcc c gene (AAZ22885) and a non-reverting an immune response against a c variety of bacteria in a mammalian host. Such vaccines can provide protection against e.g., E. coli (a cause of diarrhoea in humans), a shunonella typhimurium (the cause of mannellosis in several animal species), S. typhi (the cause of human typhiod), S. enteritidis (a cause of salmonellosis of food poisoning in humans). S. choleraesuis (a cause of salmonellosis of malingitis), S. dublin (a cause of both a systemic and diarrhoea) (arcs contile, especially of new-born calves), Haemophilus influenzae (a cause of cause of cause of cause of cause of cause of meningitis). Nelsseria gonorrhoeae (a cause of gonorrhoeae), Yersinia from gastroenteritis to fatal septicaemic disease, in humans ranging from gastroenteritis to fatal septicaemic disease), Bordetella pertusais (the cause of whooping cough), and Brucella abortus (a cause of abortion humans).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                      primers TT1-TT4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23 AEIXNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGFTQINSDLTGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New attenuated bacteria useful as vaccines for protecting against
                                                                                                                                                                                                                                                                                                                                 protein. The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This sequence represents the E. coli wild-type ompF
                                                                                                                                                                                                                                                                                                                                                      sequence of the ompF gene was removed via PCR using
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF
                                                                                                                                                                                                                                                                                               Disclosure; Page 63-64; 69pp; English.
                                                                                                           (PEPT-) PEPTIDE THERAPEUTICS LID
                                                                       98GB-0006449.
                                    99WO-GB00935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 100.
Matches 340; Conservative
                                                                                                                                                                                 WPI; 1999-580447/49.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            362 AA;
                                                                                                                                                                                                      N-PSDB; AAZ22887
                                                                                                                                              Chatfield SN;
                                                                     25-MAR-1998;
30-SEP-1999
                                                                                                                                                                                                                                                          Infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

RESULT 3

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This invention describes a novel method for predicting the tendency of a protein to form either: (a) an amphiphilic alpha-structure or (b) an amphiphilic beta-structure. The methods are used particularly for predicting the structure of membrane proteins such as glucose transporter proteins (GLUTS). They can be used to discern the function of proteins. They can also be used for the rational design or identification of compounds which interact with the proteins or to engineer proteins having particular structures. This sequence represents an Escherichia coll Ompf porlin which is used to illustrate the method of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVYYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Predicting tendency to form amphiphilic alpha and beta structures - using a novel algorithm to calculate values for subsequent graphical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 AEIYNKDGNKVDLYGKAVGLHYRSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EYEGFGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 EYBGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    THRETHTSGFANKTODVLLVAQYOFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT
                                                                                                                                                                                                                                              OmpF; porin; amphiphilic alpha-structure; beta-structure; GLUT; membrane protein; glucose transporter protein; function.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99.1%; Score 1792; DB 17; Length 340; 99.4%; Pred. No. 3.2e-146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Iserovich
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           analysis to predict protein structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fischbarg J,
AAW92998 standard; protein; 340 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Fig 2; 106pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         UYCO ) UNIV COLUMBIA NEW YORK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94US-0355844.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95WO-US16126.
                                                                                                                                                                                       E. coli OmpF porin protein.
                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 99.4
Matches 338; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Czegledy F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1996-300839/30.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                340 AA;
                                                                                                                                                                                                                                                                                                                                        Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                     409618957-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13-DEC-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14-DEC-1994;
                                                                                                                          19-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  30-JUN-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cheung M,
                                                             AAW92998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301
THE STATE OF THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
```

```
This sequence represents the E. coli wild-type ompc protein. The coding sequence of the ompc gene was removed via PCR using primers TT7-TT10 (AAZ22893-Z22895) to produce a non-reverting deletion mutation. The mutant ompc gene (AAZ22886) was used in the production of a bacterium attenuated by a non-reverting mutation in each of the ompc gene; the arcc gene (AAZ22883, AAZ22884), and the ompr gene (AAZ22883, AAZ22888). The mutant bacteria provide immunogenic activity with reduced virulence and thus can be used as a vaccine for raising an immune response against a variety of bacteria in a mammalian host. Such vaccines can provide protection against e.g., E. coli (a cause of diarrhoea in humans), Salmonella typhimurium (the cause of salmonellosis in several animal species), S. typhi (the cause of human typhiod), S. enteritials (a cause of food poisoning in humans), S. choleraeauis (a cause of salmonellosis in pigs). S. dublin (a cause of both a systemic and diarrhoeal disease in cattle, especially of new-born calvee), Hemophilus influences (a cause of meningitis), Nelsseria gonorrhoea (a cause of gonorrhoeae), Yershila form cause of a spectrum of disease in humans ranging from grant (the cause of a spectrum of disease) in humans ranging from gastroenteeritis to fatal septicemnic disease), and infertility in cattle and a condition known as undulant fever in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.1 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVYYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22 AEVYNKDGNKLDLYGKVDGLHYFS-----DNKDVDGDQTYMRLGFKGETQVTDQLTGYGQ 76
                                                                                                                                                    Bacteria; attenuation; deletion; mutant; vaccine; immune response; Gram negative; infection; diarrhoea; food poisoning; typhoid; salmonellosis; gonorrhoea; gastroenteritis; whooping cough.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New attenuated bacteria useful as vaccines for protecting against
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 367;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             58.1%; Score 1051; DB 20
58.9%; Pred. No. 2.5e-82
11ve 45; Mismatches 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Page 57-58; 69pp; English.
    į
                                                                                                                                                                                                                                                                                                                                                                                                                             (PEPT-) PEPTIDE THERAPEUTICS LTD.
AAY42549 standard; Protein; 367
                                                                                                                coli wild-type ompC protein.
                                                                                                                                                                                                                                                                                                                                                   99WO-GB00935
                                                                                                                                                                                                                                                                                                                                                                                       98GB-0006449
                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-580447/49.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           367 AA;
                                                                                                                                                                                                                                Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAZ22885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Chatfield SN;
                                                                                                                                                                                                                                                                      WO9949026-A1
                                                                                                                                                                                                                                                                                                                                                 25-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                       25-MAR-1998;
                                                                            20-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local 9
                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
```

58.9%; Pred.

Best Local Similarity

```
porin antigen or a Bacteroides antigen as a target antigen. The antigen, a porin antigen or a Bacteroides antigen as a target antigen. The novel method of diagnosing UC in a subject suspected of having inflammatory lowel disease (IBD) comprises: (1) obtaining a sample from the subject; (2) contacting the sample with a histone HI-like antigen, or perinuclear anti-neutrophil cytophasmic antibody (pANCA) reactive fragment, to form a complex of the histone HI-like antigen, or the pANCA-reactive fragment, and antibody to the histone HI-like antigen, or the panceroid of the complex indicates that the subject has UC. The paNCA-reactive histone HI-like antigen, are useful in the antigen, porin antigen and Bacteroides antigen are useful in the diagnosis, prevention and treatment of UC. The methods can also be used for identifying agents useful for treating UC. The present sequence represents a E. coli outer membrane protein c precursor.
     166
                   225
                                                                                                                                                      The invention provides a method for the diagnosis, prevention and treatment of ulcerative colitis (UC) using histone HI-like antigen, a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Use of histone Hi, porin or Bacteroides antigens as targets for the diagnosis, prevention and treatment of ulcerative colitis \cdot
                                                                                     YLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-E
                                                                                                                                                                                                    285 GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF
                                                                 RRSNGDGVGGSISYEYEGFGIVGAYGAADRTNLQE-AQPLGNGKKAEQWATGLKYDANNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ulcerative colitis; histone; H1-like antigen; porin antigen; Bacteroides antigen; IBD; pANCA; inflammatory bowel disease; didagnosis; perlunclear anti-neutrophil cytoplasmic antibody; outer membrane protein c precursor.
DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDT
                                                                                                                                                                                                                                                                                                                                                                                                                            E. coli outer membrane protein c precursor.
                                                                                                                                                                                                                                                                                                                       AAY34058 standard; protein; 367 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 8; Fig 11; 134pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99WO-US05492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0041889
                                                                                                                                                                                                                                                                                                                                                                                           23-NOV-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (REGC ) UNIV CALIFORNIA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1999-551215/46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Braun J, Cohavy O;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   367 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W09945955-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16-SEP-1999.
                                                                                                                                                                                                                                                                                                                                                         AAY34058;
                                                                                                                                                                                                                                      309
                                 134
                                                                 167
                                                                                                                                    226
121
                                                                                                                                                                                                                                                                                                         AAY34058
                                                                                                 윱
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                         ò
                               g
                                                                   ð
                                                                                                                                    à
                                                                                                                                                                   셤
```

Length 367;

DB 20;

58.1%; Score 1051;

Query Match

```
in a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The invention provides a method for diagnosing ulcerative colitis in a subject suspected of having inflammatory bowel disease. The method comprises reacting a patient sample with a portin antigen that is immunologically reactive with pANCA (perinuclear anti-neutrophil cytoplasmic antibodies) and detecting formation of a Ag-PANCA complex as indicative collits. The method is used to diagnose ulcerative collits, The method is used to diagnose represents a E. coli outer membrane protein c precursor.
                                                                                                WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                225
                                                                                                                                                                                                                                 253
                                                                                                                                                                                                                                                           YLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-E 284
                                                                                                                                                                                                                                                                           340
                                                                                                                                                                                                                                                                                                                            Diagnosing ulcerative colitis or susceptibility, by detecting complex formation between microbial porin antigen and perinuclear anti-neutrophil cytoplasmic autoantibodies
                   Gaps
                                              9
                                                         22 AEVYNKDGNKLDLYGKVDGLHYFS-----DNKDVDGDQTYMRLGFKGETQVTDQLTGYGQ 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      antigen; MAb;
                                           AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                     DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDT------A
                                                                                                                                                                                                       RRSNGDGVGGSISYEYEGFGIVGAYGAADRTNLQE-AQPLGNGKKAEQWATGLKYDANNI
                                                                                                                                                                                                                     GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF
                   34;
                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ulcerative colitis; inflammatory bowel disease; porin an pANCA; perinuclear anti-neutrophil cytoplasmic antibody; histone H1; outer membrane protein c precursor.
45; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        E. coli outer membrane protein c precursor.
                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                         AAY57356 standard; Protein; 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 1; Fig 10; 49pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0041889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96US-0057846.
97US-0837058.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                  212; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (REGC ) UNIV CALIFORNIA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2000-255695/22.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Braun J;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12-APR-1996;
11-APR-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US6033864-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Cohavy 0,
                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY57356;
                                                                                                                                                                                                       167
                                                                                                                                                                                                                                                           226
                                                                                                                                                                                                                                                                                                                                       309
                   Matches
                                                                                              61
                                                                                                                         11
                                                                                                                                                   121
                                                                                                                                                                            134
                                                                                                                                                                                                                                                                                                              285
                                                                                                                                                                                                                                                                                                                                                                                               AAY57356
                                                                    ద
                                                                                              ò
                                                                                                                      g
                                                                                                                                                   ò
                                                                                                                                                                                                      à
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
```

X S

ò 셤 ò 셤 δ q ò g ò g

Ŋ

```
22
                                                                                                                                                                                                                                                                                                                                                 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           371
                                                                                                                                                                                                                                                                                                                           6.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABB0821
g
                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                              윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                         æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAH84373 to AAH84499 represent Escherichia coli growth and proliferation related DNA sequences (I). AAH84500 to AAH84670 encode the E. coli growth and proliferation related proteins given in AAG99078 and AAG98830 to AA689899. (I) can be used as potential targets for the generation of new antimicrobial agents, and for identification of compounds which
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  coli growth and proliferation related protein sequence SEQ ID NO:450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Novel nucleic acids that inhibit Escherichia coli proliferation, useful for screening for homologous genes and for designing expression vectors
                                                                                                                                                                                                                                                                                                                           340
                                                                                                                                                                                                                         WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                            DTAYSDDFFVGRVGCVATYRNSNFFGLVDGLNFAVQYLGKNERDT------A 166
                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                          YLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-E 284
                                                                                            1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Escherichia coli; growth; proliferation; microbial; antimicrobial; bacterial infection; microorganism.
                                                                                                                                                                                                                                                                                                                       GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF
                                                        34;
                                 Length 367;
                                                        Indels
                                                        69;
                              58.1%; Score 1051; DB 21
58.9%; Pred. No. 2.5e-82;
ive 45; Mismatches 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 19; Page 505-506; 522pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                              ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ohlsen K, Zyskind J;
                                                                                                                                                                                                                                                                                                                                                                                                          AAG98980 standard; Protein; 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09-NOV-2000; 2000WO-US30950.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0164415
                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (ELIT-) ELITRA PHARM INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-335933/35.
                                             Similarity
           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAH84651
        367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200134810-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Forsyth RA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                          26-SEP-2001
                                             Best Local Sim
Matches 212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          17-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                   AAG98980;
          Sequence
                                Query Match
                                                                                                                                                                            121
                                                                                                                                                                                                   134
                                                                                                                                                                                                                                                                                                                        285
                                                                                                                                                                                                                                                                                                                                                309
                                                                                                                                                                                                                                                                                                                                                                                                AAG98980
                                                                                                                                                                                                                                                                                                                                                                                      RESULT
```

```
interact with the gene products of (I). In addition the expression of used the purification of the proteins, the purified proteins can be used to generate reagents and screen small molecule libraries or other candidate compound libraries for compounds that can be further developed to yield novel antimicrobial compounds. In addition, nucleic acid probes complementary to (I) that are specific for particular species of microorganisms can be used to identify particular microorganism species in clinical specimens, therefore, providing a rapid and dependable microorganisms can be used to identify particular microorganism species in clinical specimens, therefore, providing a rapid and dependable infection. Also, antibodies generated against proteins translated from mRNA transcribed from proliferation-required sequences can also be used to specific microorganisms that produce such proteins in a species specific manner. AAH84371 and AAH84670 represent sequencing primers used in the isolation of E. coli growth and proliferation of E. coli growth and proliferation of E. coli growth and prosine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  outer membrane protein C; Crohn's disease; immunoglobulin A; IgA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVE-- 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RSNGDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 377;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Escherichia coli outer membrane protein C (OmpC) precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                57.2%; Score 1035; DB 22;
56.9%; Pred. No. 6.3e-81;
iive 42; Mismatches 78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABB08211 standard; protein; 367 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  antiinflammatory; immune response.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17-MAY-2001; 2001WO-US16032.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           377 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        334 VGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : |: || || |
LGLVYQF 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200189361-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Sin
Matches 209;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABB08211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
```

us-09-490-291-9.rag

```
Braun
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY57355
ID AAY5
XX
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                              The sequence represents the E. coli outer membrane protein C (OmpC) precursor. The invention relates to a novel method for diagnosing Crohn's disease in a subject. The method comprises detecting immunoslobulin (Ig)A OmpC antibodies in the subject, where the presence of the IgA anti-OmpC antibodies indicates that the subject has Crohn's disease. The proteins of the invention have antiinflammatory activity. The method of the invention is an immunological response elicitor. The method is useful for diagnosing Crohn's disease in a subject and administering an OmpC antigen is useful for inducing tolerance in a patient with Crohn's disease.
                                                                                                       Diagnosing Crohn's disease in a subject comprises detecting the presence of immunoglobulin A anti-outer membrane protein C antibodies
                                                                                                                                                                                                                                                                                                                                                                                                                   WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                    DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDT-------A 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RRSNGDGVGGSISYEYEGFGIVGAYGAADRTNLQE-AQPLGNGKKAEQWATGLKYDANNI 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-E 284
                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                      22 AEVXNKDGNKLDLYGKVDGLHYFS----DNKDVDGDQTYMRLGFKGETQVTDQLTGYGG 76
                                                                                                                                                                                                                                                                                                                                                        1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF
                                                                                                                                                                                                                                                                                                                                  34;
                                                                                                                                                                                                                                                                                                           57.1%; Score 1033; DB 23; Length 367; 58.3%; Pred. No. 9e-81; ive 45; Mismatches 71; Indels 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ulcerative colitis; histone; H1-like antigen; porin antigen; Bacteroides antigen; IBD; pANCA; inflammatory bowel disease; diagnosis; perinuclear anti-neutrophil cytoplasmic antibody; outer membrane protein F precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E. coli outer membrane protein F precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
                       (CEDA-) CEDARS SINAI MEDICAL CENT
                                                           Sutton CL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY34057 standard; protein; 377
                                                                                                                                         Claim 3; Fig 5; 55pp; English.
19-MAY-2000; 2000US-0575061
                                                                                                                                                                                                                                                                                                                       Best Local Similarity 58.3 Matches 210; Conservative
                                  (REGC ) UNIV CALIFORNIA.
                                                         Targan SR, Braun J,
                                                                               WPI; 2002-075333/10.
                                                                                                                                                                                                                                                                                    367 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          W09945955-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY34057;
                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              285
                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY3405
å
                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
```

```
The invention provides a method for the diagnosis, prevention and treatment of ulcerative colitis (UC) using histone H1-like antigen, a portan artigen or a Bacteroldes antigen as a target antigen. The novel method of diagnosing UC in a subject suspected of having inflammatory bowel disease (IBD) comprises: (1) obtaining a sample from the subject; (2) contacting the sample with a histone H1-like antigen, or perinuclear anti-neutrophil cytoplasmic antigen, or the pANCA-reactive fragment, to complex of the histone H1-like antigen, or the pANCA-reactive fragment, and intubody to the histone H1-like antigen, and (3) detecting the presence or absence of the complex; where the presence of the complex indicates that the subject has UC. The pANCA-reactive histone H1-like antigen are useful in the antigen, porin antigen and bacteroldes antigen are useful in the antigen, porin antigen and treatment of UC. The methods can also be used for identifying agents useful for treating UC. The present sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNE------RDTAR 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVE-- 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LATMYSETRNMTPFGD---SDYAVANKTONFEVTAQYOFDFGLRPAVSFLMSKGRDLHAA 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVA 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22 AEVYNKDGNKLDLYGKVDGLHYFSDNSAK-----DGDQSYARLGFKGETQINDQLTGYQQ 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                           Use of histone H1, porin or Bacteroides antigens as targets for the diagnosis, prevention and treatment of ulcerative colitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RSNGDGVGGSISYEY-EGFGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 377;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  56.9%; Score 1029; DB 20; 56.9%; Pred. No. 2.1e-80; Live 41; Mismatches 79;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY57355 standard; Protein; 377 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 8; Fig 11; 134pp; English
   99WO-US05492
                                                                      98US-0041889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 56.9%
Matches 209; Conservative
                                                                                                                                            (REGC ) UNIV CALIFORNIA
                                                                                                                                                                                                                                                                                         WPI; 1999-551215/46.
                                                                                                                                                                                                                      Cohavy 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    377 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           334 VGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LGLVYOF 377
12-MAR-1999;
                                                                      12-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
```

```
:|:||||
LGLVYQF 377
                                                                                                                                                                                                                                                                                                                                                                                                                                         11-OCT-2001
                                                                                                                                                                  ABG28703;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                     371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28
                                                                                    RESULT 11
                                                                                                        ABG28703
                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The invention provides a method for diagnosing ulcerative colitis in a subject suspected of having inflammatory bowel disease. The method comprises reacting a patient sample with a porin antigen that is immunologically reactive with pANCA (perinuclear anti-neutrophil cytoplasmic antibodies) and detecting formation of a Ag-PANCA complex as indicative of ulcerative colitis. The method is used to diagnose ulcerative colitis or susceptibility to it. The present sequence represents a E. coli outer membrane protein F precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Diagnosing ulcerative colitis or susceptibility, by detecting complex formation between microbial porin antigen and perinuclear anti-neutrophil cytoplasmic autoantibodies -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNE-------RDTAR 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              168 RSNGDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIY 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       194 HENGDGWGLSTTYDLGMGFSAGAAYTSSDRTNDQVNHTAAGGDKADAWTAGLKYDANNIY 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----GIGDVDLVNYFEVGATYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVA 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 AEVYNKDĞNKLDLYĞKVDĞLHYFSDNSAK-----DĞDQŞYARLĞFKĞETQINDQLTĞYQQ 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                         porin antigen; MAb;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVE--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      56.9%; Score 1029; DB 21; Length 377; 56.9%; Pred. No. 2.1e-80; 1ve 41; Mismatches 79; Indels 38
                                                                                                                     Ulcerative colitis; inflammatory bowel disease; porin an
PANCA; perinuclear anti-neutrophil cytoplasmic antibody;
histone H1; outer membrane protein F precursor.
                                                                              coli outer membrane protein F precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 1; Fig 10; 49pp; English
                                                                                                                                                                                                                                                                                                                                 98US-0041889
                                                                                                                                                                                                                                                                                                                                                                        96US-0057846
97US-0837058
                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity .... hes 209; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                         (REGC ) UNIV CALIFORNIA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2000-255695/22.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Braun J;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 377 AA;
                                                                                                                                                                                                          Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      334 VGIVYQF 340
                                       13-JUN-2000
                                                                                                                                                                                                                                                                                                                                 12-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                          12-APR-1996;
                                                                                                                                                                                                                                                                                                                                                                                              11-APR-1997;
                                                                                                                                                                                                                                                 US6033864-A.
                                                                                                                                                                                                                                                                                          07-MAR-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Cohavy O,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
AAY57355;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   285
80 X CC CC CC CC X S X X B X B X X B X X B X X B X X B X X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
```

```
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polypeptide (II) sequences. (I) is useful as hybridisation probes, and gene mapping, and in recombinant production of (II): The and gene mapping, and in recombinant production of (II): The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques of restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or guantitating a polypeptide in tissue, as molecular weight markers and as (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. Contexpossible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and and conditional contexpossible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and alagnostic amino acid sequences of the invention.

Contexpossible dor genetic disorders or other traits the printed contexposition of mutations and to produce other types of the invention.

Contexpossible dor genetic disorders or other traits to assess biodiversity and to produce other types of the invention.

Contexpossible dor genetic disorders or other traits form when the printed contexposition and and activity of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess
                                                                                                                                                                                                                                                                                                                                             Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Score 972; DB 22; Length 573;
; Pred. No. 3e-75;
33; Mismatches 71; Indels 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 it ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 20; SEQ ID No 59062; 103pp; English.
                                                                                                                                                                                                                                                         Novel human diagnostic protein #28694.
    ż
ABG28703 standard; Protein; 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          53.7%;
61.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Tang YT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-0540217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-MAR-2001; 2001WO-US08631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23-AUG-2000; 2000US-0649167
                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 61.0%
Matches 188; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2001-639362/73.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Drmanac RT, Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          573 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAS92890
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200175067-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31-MAR-2000;
                                                                                                                                                                    18-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Home saptens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 biodiversity
```

us-09-490-291-9.rag

```
61
                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                       134
                                                                                                                                                                                                                                                                                                                                                           229
                                                                                                                                                                                                                                                                                                                                                                                299
                                                                                                                                                                                                                                                                                                                                                                                                     284
                                                                                                                                                                                                                            11
                                                                                                                                                                                                                                                                                            181
88888888888888
                                                                                                                                                                                    ద
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymersae chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in
                  D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                             314
                                                                                                         WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG
                                                                                             YEYEG--FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA
                                                                                                                                        TPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess
                                                                                                                                                                                                                                                                                                                                              ); chromosome mapping; gene mapping; gene therapy; forensic; supplement; medical imaging; diagnostic; genetic disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 20; SEQ ID No 55900; 103pp; English.
                                                                                                                                                                                                                                                                                                                         Novel human diagnostic protein #25532.
                                                                                                                                                                                                                                                           ABG25541 standard; Protein; 328 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tang YT;
                                                                                                                                                                                                                                                                                                                                                                                                                                             30-MAR-2001; 2001WO-US08631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000US-0540217
2000US-0649167
                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-639362/73.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Drmanac RT, Liu C,
                                                                                                                                                                                   GATYYFNK 305
                                                                                                                                                                                              ||||||||
315 GATYYFNK 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAS89728
                                                                                                                                                                                                                                                                                                                                                                                                  WO200175067-A2
                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31-MAR-2000;
23-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          biodiversity
                                                                                                                                                                                                                                                                                                                                                                                                                        11-OCT-2001.
                                                                                                                                                                                                                                                                                                    18-FEB-2002
                                                                                                                                                                                                                                                                               ABG25541;
                                                                                                                                                                                                                                                                                                                                                         food
                                                                        141
                                                                                             180
                                                                                                                   201
                                                                                                                                        238
                                                                                                                                                            261
                                                                                                                                                                                 298
       19
                             83
                                                  121
                                                                                                                                                                                                                                     RESULT 12
                                                                                                                                                                                                                                                 ABG2554]
       õ
                            g
                                                  ò
                                                                     셤
                                                                                             à
                                                                                                                 셤
                                                                                                                                       ò
                                                                                                                                                           g
                                                                                                                                                                                 ò
                                                                                                                                                                                                     셤
```

```
diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |:|||| | | : ||: | | : ||1 | 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                              1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EYEGFCIVGAYGAADRTNLQE-AQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240 ITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-EGIGDVDLVNYFEVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
                                                                                                                                                                                                                                                                                                                                45;
                                                                                                                                                                                                                                                                             Length 328;
                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                DB 22;
                                                                                                                                                                                                                                                                                                                              63;
                                                                                                                                                                                                                                                                             52.9%; Score 957.5; DB 2
56.4%; Pred. No. 2.5e-74;
ive 43; Mismatches 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Novel human diagnostic protein #25859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABG25868 standard; Protein; 759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-MAR-2001; 2001WO-US08631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000US-0540217
2000US-0649167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tang
                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 56.4
Matches 195; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2001-639362/73.
N-PSDB; AAS90055.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Liu C,
                                                                                                                                                                                                                                328 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200175067-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31-MAR-2000;
23-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11-0CT-2001,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Drmanac RT,
                                                                                                                                                                                                                                  Sequence
```

```
WO200175067-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 299
                                                                                                                                                                                                                                                                                                                                                                                                              61
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                  .;
6
                                                                                                                                                                     Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                           WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                   DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                                                                                                                                                    181 EYEGFGIVGAYGAADRINLQE-AQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATP 239
                                                                                                                                                                                                                                                                                                                                             240 ITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-EGIGDVDLVNYFEVG 298
                                                                                                                                                                                                                                                                                                                                                                         453 AEVYNKDGNKLDLYGKVDGLHYFS-----DNKDVDGDQTYMRLGFKGETQVTDQLTGYGQ 507
                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                               1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
responsible for genetic disorders or other traits and to assess
                                                                                                                                                                                                                                 45;
                                                                                                                                                                                                                   DB 22; Length 759;
                                                                                                                                                                                                                        ,7e-74;
.es 63; Indels
                                                                                                                                                                                                                                                                                                                                                                                                       299 ATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                 Query Match 52.9%; Score 957.5;
Best Local Similarity 56.4%; Pred. No. 7.7e
Matches 195; Conservative 43; Mismatches
                     Claim 20; SEQ ID No 56227; 103pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Novel human diagnostic protein #29784.
                                                                                                                                                                                                                                                                                                                                                                                                                                                ABG29793 standard; Protein; 759 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                    759 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-FEB-2002
                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABG29793;
                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                          ABG29793
                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                        g
õ
                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                 요
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, colymerase chair reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags contentifying expressed genes. (I) is useful in gene therapy techniques contentifying expressed genes. (I) is useful for generating antibodies against it, detecting or quantifating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynuclectide sequences have applications in capposition for genetic disorders or other trails to assess blodiversity and to produce other types of data and products dependent on DNA and canno acid sequences. ABG00010-ABG30377 represent novel human call sequence data for this patent did not appear in the printed considers in the printed security in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DIAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          660 V----GSLGWANKAQNFEAVAQYQFDFGLRPSLAYLQSKGKNLGRGYDDEDILKYVDVG 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       508 WEYQIQGNSAEN---ENNSWTRVAFAGLKFQDVGSFDYGRNYGVYYDVTSWTDVLPEFGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18.1 EYEGFGIVGAYGAADRTNLQE-AQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240 ITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-EGIGDVDLVNYFEVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 22; Length 759;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               52.9%; Score 957.5; DB 2
56.4%; Pred. No. 7.7e-74;
tive 43; Mismatches 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              at ftp.wipo.int/pub/published_pct_sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 20; SEQ ID No 60152; 103pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Tang YT;
                                                                                           30-MAR-2001; 2001WO-US08631.
                                                                                                                                                                                                       31-MAR-2000; 2000US-0540217
23-AUG-2000; 2000US-0649167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Drmanac RT, Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2001-639362/73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 195; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               759 AA;
                                                                                                                                                                                                                                                                                                                                                             (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAS93980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             biodiversity
11-0CT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
```

8

```
715 ATYYFNKNMSTYVDYKINLLD-DNQFTRDAGINTDNIVALGLVYQF 759
```

qq

```
8;
                                                                                                                                                                                                                                                                                                                                                                                                                                         A perchloroethylene dehalogenase derived from Clostridium bifermentans DPH-1, for decomposing chlorinated aliphatic compounds. -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence represents perchloroethylene (PCE) dehalogenase derived from Clostridium bifermentans DPH-1. PCE dehalogenase: (a) catalyses the dehalogenising reaction forming trichloroethylene from tetrachloroethylene; (b) is a dimer consisting of two subunits having a molecular weight of 35 kDa; (c) has an optimum pt of 7.5; (e) has enzymatic activity that is not affected by the addition of metalions, ethylene diamine tetraacetate, reduced nicotinamide adenine dinucleotide or vapanocobalami; and (f) has enzymatic activity which decomposing chlorinated aliphatic compounds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDT------ 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ARRSNGDGVGGSISYEYEGFGIV--GAYGAADRINLQEAQPLGNGKKAEQWATGLKYDAN 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NIYLAANYGETRNAIPIINKFINTSGFANKIQDVLLVAQYQFDFGLRPSIAYTKSKAKDV 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                        C. bifermentans DPH-1 perchloroethylene dehalogenase SEQ ID NO:1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        85; Indels 30;
                                                                                                                                                     Clostridium bifermentans DPH-1; perchloroethylene dehalogenase; enzyme; decomposing; chlorinated aliphatic compound.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 366;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        52.3%; Score 946; DB 23;
55.8%; Pred. No. 2.8e-73;
ive 36; Mismatches 85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 4; Page 11-12; 15pp; Japanese.
                             ABB06345 standard; Protein; 366 AA
                                                                                                                                                                                                                                                                                              04-JUL-2000; 2000JP-0202729
                                                                                                                                                                                                                                                                                                                               04-JUL-2000; 2000JP-0202729
                                                                                         30-MAY-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 55.8
Matches 191; Conservative
                                                                                                                                                                                                     Clostridium bifermentans.
                                                                                                                                                                                                                                                                                                                                                              (GIFU-) GIFU DAIGAKUCHO
                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2002-263241/31.
N-PSDB; ABL49767.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            366 AA;
                                                                                                                                                                                                                                  JP2002017358-A
                                                                                                                                                                                                                                                                 22-JAN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seguence
                                                            ABB06345;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
RESULT 15
ABB06345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      254
                                                            g
```

ð

ò g ò 셤 ð g ð

```
284 EGI----GDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSD 321
ò
                g
```

Search completed: December 18, 2002, 16:17:26 Job time: 24.7682 secs

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 340 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS: single
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Y: USA
10112-0228
                                1000000000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 1
US-08-355-844-1
                                                 108.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
1088.5
108
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Appl
Appl
Appli
Appli
                                                                                                                                                                                     (without alignments)
1277.936 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                            December 18, 2002, 16:14:14; Search time 7.85111 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence Sequence Sequence S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence Seq
                                                                                                                                                                                                                                                     US-09-490-291-9
1809
1 AEIYNKDGNKVDLYGKAVGL.....NKLGVGSDDTVAVGIVYQFA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Issued_Patents_AA:*
1: /cgn2_6/ptodatca1/laa/5A_COMB.pep:*
2: /cgn2_6/ptodatca1/laa/5B_COMB.pep:*
3: /cgn2_6/ptodatca1/laa/6A_COMB.pep:*
4: /cgn2_6/ptodatca1/laa/6B_COMB.pep:*
5: /cgn2_6/ptodatca1/laa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodatca1/laa/PcTUS_COMB.pep:*
                   GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PCT-US95-16126-1

US-09-041-889-29

US-09-041-889-29

US-09-041-889-29

US-08-09-182A-4

US-08-877-109-4

US-08-877-109-4

US-08-877-109-6

US-08-877-109-6

US-08-877-109-6

US-08-97-109-6

US-08-97-109-6

US-08-09-182A-2

US-08-09-182A-2

US-08-09-181A-12

US-08-09-181A-14

US-08-09-181A-14

US-08-09-181A-14

US-08-09-181A-14

US-08-09-181A-14

US-08-09-181A-14

US-08-09-181A-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                              262574 seqs, 29422922 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                  protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein
                                                                                                                                                                                                                                                                                                            Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database
                                                                                                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ĕ
```

PCT-US94-08326-8 US-08-096-181A-10 PCT-US94-08326-10 US-09-333-611-14 US-09-333-611-12

```
Sequence 20, A Sequence 20, A Sequence 20, A Sequence 21, A Sequence 21, A Sequence 21, A Sequence 21, A Sequence 30, A Sequence 10, A Sequen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/08355844

Patent No. 5940307

GENERAL INFORMATION:

APPLICANT: Czegledy, Ferenc

APPLICANT: Li, Jun

APPLICANT: Li, Jun

APPLICANT: Cheung, Min

TITLE OF INVENTION: STRUCTURE

NUMBER OF SEQUENCES:

CORRESCONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPOURE: IEM PC COMPAILS
COMPOURE: IEM PC COMPAILS
SOUTHARE: PACENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/355,844
FILING DATE: 14-DEC-1994
FLING DATE: 14-DEC-1994
CLASSFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: 29,705
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION 169-256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond STREET: 30 Rockefeller Plaza CITY: New York STATE: NY COUNTRY: USA
US-09-556-877-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                porin protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OmpF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: peptide ORIGINAL SOURCE: ORGANISM: Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ) NAME/KEY: Peptide
) LOCATION: 1..340
) OTHER INFORMATION:
US-08-355-844-1
```

```
g
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                    ö
                                                                                                                                               DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                              EYEGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI 240
                                                                                                                                                                                                                                                                                                                      TNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT 300
                                                                                                                                                                                                                                                                                                                                       WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                    Gaps
                                                                                  1 AEIYNKDGNKVDLYGKAVGLHYRSKGNGENSYGGNGDMIYARLGFKGETQINSDLTGYGG 60
                                                                                                                                                                                                                                                                          1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                    ö
    Length 340;
                                    1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1, Application PC/TUS9516126
GENERAL INFORMATION:
APPLICANT: Flachbarg, Jorge
APPLICANT: Caegledy, Ferenc
APPLICANT: Lay Jun
APPLICANT: Li, Jun
APPLICANT: Cheung, Min
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
TITLE OF INVENTION: STRUCTURE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      E: Brumbaugh, Graves, Donohue & Raymond
30 Rockefeller Plaza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPOUR: IBM PC compatible
COMPOUR: IBM PC compatible
OPERATING SYSTEM: PC-DOS/NS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURREMY APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16126
                                                                                                                                                                                                                                                                                                                                                                                   YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                      Score 1792; DB 2;
Pred. No. 3e-159;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A29927-50/29910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/355,844
FILING DATE: 14-DEC-1994
ATTOCNEY AGENT INFORMATION:
NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A29927-50/;
TELECOMMUNICATION:
TELEPHONE: 212-408-2586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
 99.1%;
99.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 340 amino acids TYPE: amino acid
                                 Matches 338; Conservative
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino a STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PCT-US95-16126-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET:
                                                                                                                               61
                                                                                                                                                                                                                                                                                                                        241
                                                                                                                                                                                            121
                                                                                                                                                                                                                                                         181
                                                                                                                                                                                         õ
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                      ç
                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                ð
                                                                                              g
                                                                                                                               à
                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                   ð
```

```
Sequence 30, Application US/09041889
Patent No. 6033864
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Cohavy, Offer
TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
TITLE OF INVENTION: Microbial UC pANCA antigens
NUMBER OF SEQUENCES: 4
CORRESPONDENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                 61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVYYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                      DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                             1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                         EYEGFGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                Score 1792; DB 5;
Pred. No. 3e-159;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF
                                                                                                                              OmpF porin protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-PM 3006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US/09/041,88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                Query Match
Best Local Similarity 99.1%;
Matches 338; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE: 11-APR-1997
ATTORNEY/AGENT INFORMATION:
                                                   ORGANISM: Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA
APPLICATION NUMBER: U
               MOLECULE TYPE: peptide ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                     ) OTHER INFORMATION:
PCT-US95-16126-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: San Diego
STATE: California
COUNTRY: USA
linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                        NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-041-889-30
                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
```

ö

```
Diagnosis, Prevention and Treatment of Ulcerative Colitis, and Clinical Subtypes Thereof, Using Microbial UC pANCA antigens
                                                                                                                                                                                                                                                                                                                                                                           168 RSNGDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIY 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVA 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |: | || | | : | || || 311 GGADNPAGVDDKDLVKYADIGATYFNKNMSTYVDYKINLIDEDDSFYAANGISTDDIVA 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                      DI-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNE------RDTAR 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVE-- 284
                                                                                                                                                                                                                                                                                                     1. AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGO 60
                                                                                                                                                                                                                                                                                                                       22 AEVYNKDGNKLDLYGKVDGLHYFSDNSAK-----DGDQSYARLGFKGETQINDQLTGYQQ 76
                                                                                                                                                                                                                                                                     38;
                                                                                                                                                                                                                                Length 377;
                                                                             .
                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: PATENTEN PC-DOS/NS-DOS SOFTWARE: PATENTIN Release #1.0, Version #1.25 CURREW APPLICATION DATA: APPLICATION NUMBER: US/09/041,889
                                                                                                                                                                                                                            56.9%; Score 1029; DB 3; 56.9%; Pred. No. 4.8e-88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          E: Campbell & Flores LLP
4370 La Jolla Village Drive, Suite 700
                                                                                                                                                                                                                                                                 41; Mismatches
     REFERENCE/DOCKET NUMBER: P-PM 3006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 28, Application US/09041889;
Patent No. 6033864
GENERAL INFORMATION:
APPLICANT: Braun, Jonathan
APPLICANT: Cohavy, Offer
TITLE OF INVENTION: Diagnosis, Prev
TITLE OF INVENTION: Ulcerative Coli
TITLE OF INVENTION: Microbial UC pA
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell 6 Flores LLP
                   TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-901
TELEPAX: (619) 535-9049
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                        peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRECT: 43/v __
STREET: 5an Diego
CITY: San Diego
                                                                                                                                                          Linear
                                                                                                                                                                                                                                             Best Local Similarity
Matches 209; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     334 VGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          371 LGLVYQF 377
                                                                                                                                                                  , MOLECULE TYPE:
US-09-041-889-29
                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 5
US-09-041-889-28
                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             285
                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 29, Application US/09041889
Sequence 29, Application US/09041889
Patent No. 6033864
GENERAL INFORMATION:
APPLICANT: Braun, Jonathan
APPLICANT: Cohavy, Offer
TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using TITLE OF INVENTION: Microbial UC PANCA antigens
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
                                                                                                                                                                                                                                               ..
8
                                                                                                                                                                                                                                                                                                                                                    DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDT-------A 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                 226 YLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-E 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        285 GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                   22 AEVYNKDGNKLDLYGKVDGLHYFS-----DNKDVDGDQTYMRLGFKGETQVTDQLIGYGQ 76
                                                                                                                                                                                                                                                                                1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     167 RRSNGDGVGGSISYEYEGFGIVGAYGAADRTNLQE-AQPLGNGKKAEQWATGLKYDANNI
                                                                                                                                                                                                           Length 367;
                                                                                                                                                                                                                                           69; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ZIP: 92122

ZIP: 92122

MEDIUM TYPEN FLORM:
MEDIUM TYPEN FLORPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .....yens
.....resSE: .....resS:
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
                                                                                                                                                                                                       58.1%; Score 1051; DB 3;
58.9%; Pred. No. 4.1e-90;
iive 45; Mismatches 69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/837,058
FILING DATE: 11.APR-1997
ATTORNEY/AGENT INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: Campbell, Cathryn A. REGISTRATION NUMBER: 31,815
             TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
                                                                                        LENGTH: 367 amino acids TYPE: amino acid
                                                                                                                                                                                                       Query Match 58.1%
Best Local Similarity 58.9%
Matches 212; Conservative
                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-041-889-29
                                                                                                                                                                    US-09-041-889-30
                                                                                                                TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

```
Sequence 4, Application US/08877109
Patent No. 5747287
GENERAL INFORMATION:
APPLICANT: Blake, Milan S.
APPLICANT: Tai, Joseph Y.
APPLICANT: Tai, Joseph Y.
APPLICANT: Hronowski Ludan J.J.
APPLICANT: Hronowski Ludan J.J.
APPLICANT: Hronowski Ludan J.J.
APPLICANT: Hronowski Ludan J.J.
APPLICANT: Pullen, Jeffrey K.
TITLE OF INVENTION: Method for the High Level Expression,
TITLE OF INVENTION: Purification and Refolding of the Outer Membrane Group B.
TITLE OF INVENTION: Porin Proteins from Neisseria meningitidis
CORRESPONDENCE ADDRESS:
ADDRESSED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGR--NYGVVYDALGYTDMLPEFG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120 GDTAYSDDFFVGRVGGVATYRNSNFFG--LVDGLNFAVQYLGK---NERDTARRSNGDGV 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :| | :| | : | : | : | : | 114 SIT -- EDVLGLGTIGRVESREISVRYDSPVFAGFSGSVQYVPRDNANDVDKYKHTKSSRE 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               175 GGSISYEYEGFGIVGAYGA-----ADR--TNLQEAQPLGNGKKAEQWATGLKYD 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              172 SYHAGLKYENAGFFGQYAGSFAKYADLNTDAERVAVNTANAHPV---KDYQVHRVVAGYD 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          222 ANNIY--LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSK 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               229 ANDLYVSVAGOYEAAKN----NEVGSTKGKKHEQTQVAATAAYRFG-NVTPRVSYAHGF 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   58; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 VTLYGTIKAGVEVSRVKDAGTYKAQGGKSKTATQIADFGSKIGFKGQEDLGNGMKAIWQL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11 VDLYGKAVGLHYFSKGNGENSYGGNGDMTY-----ARLGFKGETQINSDLTGYGQW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 8.7%; Score 157; DB 1; Length 342; Best Local Similarity 24.7%; Pred. No. 8.5e-07; Matches 83; Conservative 45; Mismatches 150; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sterne, Kessler, Goldstein & Fox 10 New York Ave., NW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRALLATION.
APPLICATION.
FILING DATE: 23-0...
FILING DATE: 435
ATTORNEY AGENT INFORMATION:
NAME: Esmond, RObert W.
REGISTRATION NUMBER: 32,893
REPERBONE/DOCKET NUMBER: 1438.0060000
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acids
TYPE: amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              280 AKDVEGIGDVDLVNY----FEVGATYYFNKNMSTYV 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: 1100 New York Ave., CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE: D.C.
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-877-109-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ф
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 4, Application US/08096182A
Patent No. 5439808
GENERAL INFORMATION:
APPLICANT: Blake, Milan S.
APPLICANT: Tai, Joseph Y.
APPLICANT: Qi, Huilin L.
APPLICANT: Liang, Shu-Mei
APPLICANT: Liang, Shu-Mei
APPLICANT: Pullen, Jeffrey K.
TITLE OF INVENTION: Method for the High Level Expression,
APPLICANT: Pullen, Pulfication and Refolding of the Outer Membrane Group F
TITLE OF INVENTION: Purification and Refolding of the Outer Membrane Group F
TITLE OF INVENTION: Porin Proteins from Neisseria Meningitidis
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RSNGDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIY 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGI 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                  25;
                                                                                                                                                                                                                                                                                                                                                                                    Query Match 47.5%; Score 858.5; DB 3; Length 323; Best Local Similarity 57.2%; Pred. No. 3.1e-72; Matches 174; Conservative 33; Mismatches 72; Indels 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: The Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3: Sterne, Kessler, Goldstein & Fox 1100 New York Ave., Suite 600
                                                            NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFRENCE/CDCKET NUMBER: 9-PM 3006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEPHONE: (619) 535-9001
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 323 anino acids
TYPE: amino acid
TOPOLOGY: linear
APPLICATION NUMBER: US 08/837,058 FILING DATE: 11-APR-1997 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                ; MOLECULE TYPE: peptide US-09-041-889-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY: Washington STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GDVD 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         311 GGAD 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-096-182A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
```

```
Beltsville, Maryland 20705
United States of America
Tai, Joseph Y.
Qi, Huilin L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The Rockefeller University
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1230 York Avenue
New York, New York 10021
United States of America
12103 Indian Creek Court
                                                                                                            COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Liang, Shu-Mei
Hronowski, Lucjan J.J.
Pullen, Jeffrey K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 4, Application PC/TUS9408327 GENERAL INFORMATION:
                                                        ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; MOLECULE TYPE: protein US-08-798-760-4
  Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PCT-US94-08327-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 4, Application US/08798760

Patent No. 6013267

GENERAL INFORMATION:

GENERAL INFORMATION:

APPLICANT: Blake, Milan S.

APPLICANT: Tai, Joseph Y.

APPLICANT: Liang, Shu-Mei

APPLICANT: Liang, Shu-Mei

APPLICANT: Hronowski, Lucjan J.J.

APPLICANT: Hronowski, Lucjan J.J.

APPLICANT: Hronicant Method for the High Level Expression,

TITLE OF INVENTION: Purification and Refolding of the Outer Membrane Group in TITLE OF INVENTION: Porin Proteins from Neisseria Meningitidis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 EYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGR--NYGVVYDALGYTDMLPEFG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63 E---QKASIAGINSGWGN--RQSFIGLK----GGFGTVRAGNLNTVLKDSGDNVNAWESG 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               229 ANDLYVSVAGQYEAAKN-----NEVGSTKGKKHEQTQVAATAAYRFG-NYTPRVSYAHGF 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120 GDTAYSDDFFVGRVGGVATYRNSNFFG--LVDGLNFAVQYLGK---NERDTARRSNGDGV 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 175 GGSISYEYEGFGIVGAYGA------ADR--TNLQEAQPLGNGKKAEQWATGLKYD 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           222 ANNIY--LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSK 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 VDLYGKAVGLHYFSKGNGENSYGGNGDMTY-----ARLGFKGETQINSDLTGYGQW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8.7%; Score 157; DB 1; Length 342; 24.7%; Pred. No. 8.5e-07;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
              SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             45; Mismatches 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sterne, Kessler, Goldstein & Fox 1100 New York Ave., Suite 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                283 KAKVNGVKD---ANYQYDQVIVGADYDFSKRTSALV 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        280 AKDVEGIGDVDLVNY----FEVGATYYFNKNMSTYV 311
                                                                                    CLASSIFICATION: 435
PRIOR APPLICATION 1435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,264
FILING DATE: 28-APR-1995
APPLICATION NUMBER: 08/096,182
FILING DATE: 3-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 1438.0060001
TELECOMMUNICATION INFORMATION:
TELEFHONE: (202) 371-2600
TELEFAX: (202) 371-2600
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
TYPE: amino acid
TYPE: allinear
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE: protein US-08-877-109-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, K
STREET: 1100 New Yor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
les 83; Conservat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 8
US-08-798-760-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Loca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
```

```
63 E---QKASIAGTNSGWGN--RQSFIGLK----GGFGTVRAGNLNTVLKDSGDNVNAWESG 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        229 ANDLYVSVAGQYEAAKN----NEVGSTKGKKHEQTQVAATAAYRFG-NVTPRVSYAHGF 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62 EYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGR--NYGVVYDALGYTDMLPEFG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120 GDTAYSDDFFVGRVGGVATYRNSNFFG--LVDGLNFAVQYLGK---NERDTARRSNGDGV 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 SNT--EDVLGLGTIGRVESREISVRYDSPVFAGFSGSVQYVPRDNANDVDKYKHTKSSRE 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            222 ANNIY--LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSK 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             175 GGSISYEYEGFGIVGAYGA-----ADR--TNLQEAQPLGNGKKAEQWATGLKYD 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 VTLYGTIKAGVEVSRVKDAGTYKAQGGKSKTATQIADFGSKIGFKGQEDLGNGMKAIWQL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11 VDLYGKAVGLHYFSKGNGENSYGGNGDMTY-----ARLGFKGETQINSDLTGYGQW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 342;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
CURRENT APPLICATION PRESS #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/798,760
ETLING DATE: 11-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 1438.0060002
TELEFORMUNICATION INFORMATION:
TELEFORMUNICATION INFORMATION:
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
8.7%; Score 157; DB 3; L.
Best Local Similarity 24.7%; Pred. No. 8.5e-07;
Matches 83; Conservative 45; Mismatches 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   283 KAKVNGVKD---ANYQYDQVIVGADYDFSKRISALV 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              280 AKDVEGIGDVDLVNY----FEVGATYYFNKNMSTYV 311
```

```
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ) TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-096-182A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 83; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
US-08-877-109-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16;
                      Expression, Purification and Refolding of the Outer Membrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63 E---QKASIAGTINSGWGN--RQSFIGLK----GGFGTVRAGNLNTVLKDSGDNVNAWESG 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62 EYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGR--NYGVVYDALGYTDMLPEFG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GDTAYSDDFFVGRVGGVATYRNSNFFG--LVDGLNFAVQYLGK---NERDTARRSNGDGV 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            175 GGSISYEYEGFGIVGAYGA------ADR--TNLQEAQPLGNGKKAEQWATGLKYD 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                172 SYHAGLKYENAGFFGOYAGSFAKYADLNTDAERVAVNTANAHPV---KDYQVHRVVAGYD 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ANNIY -- LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSK 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 VDLYGKAVGLHYFSKGNGENSYGGNGDMTY-----ARLGFKGETQINSDLTGYGQW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Group B
Porin Proteins from Neisseria Meningitidis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8.7%; Score 157; DB 5; Length 342; 24.7%; Pred. No. 8.5e-07; ative 45; Mismatches 150; Indels
                                                                                                                                                                                                                ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/08327
FILING DATE: Herewith
CLASSIFICATION:
                                                                                                                          SEE: Sterne, Kessler, Goldstein & Fox : 1100 New York Ave., Suite 600 Washington
     Method for the High Level
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  280 AKDVEGIGDVDLVNY----FEVGATYYFNKNMSTYV 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Esmond, Robert W.
REGISTRATION UNDBER: 32,893
REFERENCE/DOOKET NUMBER: 1438.006PC00
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                     US 08/096,182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
US-08-096-182A-6
Sequence 6, Application US/08096182A
Patent No. 5439808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEG ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08
FILING DATE: 23 July 1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 24.79
nes 83; Conservative
TITLE OF INVENTION: Meth
TITLE OF INVENTION: EXPI
TITLE OF INVENTION: PUTI
TITLE OF INVENTION: POTI
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Res
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PCT-US94-08327-4
                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                 STREET:
                                                                                                                                                                                  STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
```

```
APPLICANT: Blake, Milan S.
APPLICANT: Tai, Joseph Y.
APPLICANT: Tai, Joseph Y.
APPLICANT: Liang, Shu-Mei
APPLICANT: Hronowski, Lucjan J.J.
APP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :| ; | ; | ; | ; | ; | ; | 192
135 SNT--EDVLGLGTIGRVESREISVRYDSPVFAGFSGSVQYVPRDNANDVDKYKHTKSSRE 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 EYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGR--NYGVVYDALGYTDMLPEFG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        222 ANNIY -- LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSK 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ||::| :| ;| :| | :| | :| | :| | 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120 GDTAYSDDFFVGRVGGVATYRNSNFFG--LVDGLNFAVQYLGK---NERDTARRSNGDGV 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     175 GGSISYEYEGFGIVGAYGA------ADR--TNLQEAQPLGNGKKAEQWATGLKYD 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24 VILYGTIKAGVEVSRVKDAGTYKAQGGKSKTATQIADFGSKIGFKGQEDLGNGMKAIWQL 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11 VDLYGKAVGLHYFSKGNGENSYGGNGDMTY-----ARLGFKGETQINSDLTGYGQW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8.7%; Score 157; DB 1; Length 363; 24.7%; Pred. No. 9.2e-07; 1ve 45; Mismatches 150; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/096,182A
FILING DATE: 23-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ESMONG, ROBERT W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
CADDRESSEE: Sterne, Kessler, Goldstein & Fox STREET: 1100 New York Ave., Suite 600
CITY: Washington
CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     280 AKDVEGIGDVDLVNY----FEVGATYYFNKNMSTYV 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 1438
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : 363 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
```

```
APPLICANT: Blake, Milan S.
APPLICANT: Tai, Joseph Y.
APPLICANT: Tai, Joseph Y.
APPLICANT: Tai, Joseph Y.
APPLICANT: Liang, Shu-mei
APPLICANT: Liang, Shu-mei
APPLICANT: Hronowski, Lucjan J.J.
APPLICANT: Hronowski, Jeffichy K.
TITLE OF INVENTION: Portin Proteins from Neisseria Meningitidis
NUMBER OF SEQUENCE: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :| | :| | : | : | : | : | 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 193 SYHAGLKYENAGFFGQYAGSFAKYADLNTDAERVAVNTANAHPV---KDYQVHRVVAGYD 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ANNIY--LAANYGETRNATPITNKETNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSK 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62 EYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGR--NYGVVYDALGYTDMLPEFG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120 GDTAYSDDFFVGRVGGVATYRNSNFFG--LVDGLNFAVQYLGK---NERDTARRSNGDGV 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .75 GGSISYEYEGFGIVGAYGA-----ADR--TNLQEAQPLGNGKKAEQWATGLKYD 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84 E---OKASIAGINSGWGN--ROSFIGLK----GGFGTVRAGNLNTVLKDSGDNVNAWESG 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 VDLYGKAVGLHYFSKGNGENSYGGNGDMTY-----ARLGFKGETQINSDLTGYGQW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24 VILYGTIKAGVEVSRVKDAGTYKAQGGKSKTATQIADFGSKIGFKGQEDLGNGMKAIWQL 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 8.7%; Score 157; DB 3; Length 363; Best Local Similarity 24.7%; Pred. No. 9.2e-07; Matches 83; Conservative 45; Mismatches 150; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE: D.C.
COUNTRY: USA
COUNTRY: USA
ZIP: 20005
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: PRESTICE PACENTER PRESTICE PACENTER PRESTICE PACENT RELEASE #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/798,760
FILING PATE: 11-FEB-1997
FILING PATE: 11-FEB-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Ave., Suite 600
                                              ATTORNEY AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOOKET NUMBER: 1438.0060002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2600
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
           280 AKDVEGIGDVDLVNY ----FEVGATYYFNKNMSTYV
                                                                                                                                                                                                     Sequence 6, Application US/08798760 Patent No. 6013267 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1: 363 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; TOPOLOGY: Ilnear
; MOLECULE TYPE: protein
US-08-798-760-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                222
              ö
                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
Sequence 6, Application US/08877109
Patent No. 5747287
GENERAL INFORMATION:
APPLICANT: Blake, Milan S.
APPLICANT: Tai, Joseph Y.
APPLICANT: Liang, Shu-Mei
APPLICANT: Liang, Shu-Mei
APPLICANT: Hydnowski, Lucjan J.J.
APPLICANT: Pullen, Befolding of the Outer Membrane Group ITILE OF INVENTION: Purification and Refolding of the Outer Membrane Group ITILE OF INVENTION: Porin Proteins from Neisseria meningitidis
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 EYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGR--NYGVVYDALGYTDMLPEFG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120 GDTAYSDDFFVGRVGGVATYRNSNFFG--LVDGLNFAVQYLGK---NERDTARRSNGDGV 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        175 GGSISYEYEGFGIVGAYGA-----ADR--TNLQEAQPLGNGKKAEQWATGLKYD 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      222 ANNIY - - LAANYGETRNATPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSK 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24 VTLYGTIKAGVEVSRVKDAGTYKAQGGKSKTATQIADFGSKIGFKGQEDLGNGMKAIWQL 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11 VDLYGKAVGLHYFSKGNCENSYGGNGDMTY-----ARLGFKGETQINSDLTGYGQW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 363;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 8.7%; Score 157; DB 1; Length 36 Best Local Similarity 24.7%; Pred. No. 9.2e-07; Matches 83; Conservative 45; Mismatches 150; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,109
                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: Sterne, Kessler, Goldstein & Fox STREET: 1100 New York Ave., NW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE/DOCKET NUMBER: 1438.0060001
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,264
FILING DATE: 28-APR-1995
APPLICATION NUMBER: 08/096,182
FILING DATE: 23-JUL-1993
ATTORNEY AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REGISTRATION NUMBER: 32,893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : 363 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE: protein US-08-877-109-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
```

엄 à 셤 ò . В à 엄 à

```
APPLICANT: Blake, Milan S.
APPLICANT: Tai, Joseph Y.
APPLICANT: Tai, Joseph Y.
APPLICANT: Tai, Joseph Y.
APPLICANT: Liang, Shu-Mei
APPLICANT: Liang, Shu-Mei
APPLICANT: Liang, Shu-Mei
APPLICANT: Hronowski, Lucjan J.J.
APPLICANT: Pullen, Jeffrey K.
TITLE OF INVENTION: Method for the High Level Expression,
TITLE OF INVENTION: Porification and Refolding of the Outer Membrane Group F
TITLE OF INVENTION: Porin Proteins from Neisseria Meningitidis
CORRESPONDENCES: 23
       84 E---OKASIAGINSGWGN--RQSFIGLK----GGFGTVRAGNLNTVLKDSGDNVNAWESG 134
                                                                                     193 SYHAGLKYENAGFFGQYAGSFAKYADLNTDAERVAVNTANAHPV---KDYQVHRVVAGYD 249
                                                                                                                                                                                                                                                 222 ANNIY--LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSK 279
                                                                                                                                                                                                                                                                                              250 ANDLYVSVAGQYEAAKN----NEVGSTKGKKHEQTQVAATAAYRFG-NVTPRVSYAHGF 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      35 NGDMTY------ARLGFKGETQINSDLTGYGQWEYNFQGNNSEGADAQTGNKTRL 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19 NGQVTEVTTATGIVDLGSKIGFKGQEDLGNGLKAIWQVE---QKASIAGTDSGWGN--RQ 73
                                                    120 GDTAYSDDFFVGRVGGVATYRNSNFFG--LVDGLNFAVQYLGK---NERDTARRSNGDGV
                                                                                                                                                    175 GGSISYEYEGFGIVGAYGA-----ADR--TNLQEAQPLGNGKKAEQWATGLKYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 1; Length 309;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
8.2%; Score 147.5; DB 1;
Best Local Similarity 25.1%; Pred. No. 5.7e-06;
Matches 76; Conservative 46; Mismatches 116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEE: Sterne, Kessler, Goldstein & Fox : 1100 New York Ave., Suite 600 Washington
                                                                                                                                                                                                                                                                                                                                                280 AKDVEGIGDVDLVNY----FEVGATYYFNKNMSTYV 311
                                                                                                                                                                                                                                                                                                                                                                                           304 KAKVNGVKD---ANYQYDQVIVGADYDFSKRISALV 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1438.0060000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/08/096,182A FILING DATE: 23-JUL-1993 CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZIP: 2005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Pales
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 2, Application US/08096182A Patent No. 5439808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY AGENT INFORMATION:
NAME: ESENONG, ROBERT W.
REGISTATION NUMBER: 32,893
REPERENCE/DOCKET NUMBER: 1438.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE GHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
STREET: 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
US-08-096-182A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: Was
STATE: D
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-096-182A-2
                                                                                                 ద
                                                                                                                                                                                              g
                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                        g
                                                       ò
                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Expression, Purification and Refolding of the Outer Membrane
250 ANDLYVSVAGQYEAAKN----NEVGSTKGKKHEQTQVAATAAYRFG-NVTPRVSYAHGF 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGR--NYGVVYDALGYTDMLPEFG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11 VDLYGKAVGLHYFSKGNGENSYGGNGDMTY-----ARLGFKGETQINSDLTGYGQW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Q1, Hullin L.
APPLICANT: Liang, Shu-Mei
APPLICANT: Liang, Shu-Mei
APPLICANT: Liang, Shu-Mei
APPLICANT: Hronowski, Lucjan J.J.
APPLICANT: Hronowski, Lucjan J.J.
APPLICANT: Hronowski, Lucjan J.J.
APPLICANT: Pullen, Jeffrey K.
TITLE OF INVENTION: Expression,
TITLE OF INVENTION: Purification and Refolding of the Outer Mem
TITLE OF INVENTION: Group B
TITLE OF INVENTION: Porin Proteins from Neisseria Meningitidis
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 363;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/08327
FILING DATE: HERWITH
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N 8.7%; Score 157; DB 5; I Similarity 24.7%; Pred. No. 9.2e-07; 33; Conservative 45; Mismatches 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                   Tai, Joseph Y.
Oi, Hullin L.
Liang, Shu-Mei
Hronowski, Lucjan J.J.
Pullen, Jeffrey K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3: Sterne, Kessler, Goldstein & Fox 1100 New York Ave., Suite 600
                                                                             280 AKDVEGIGDVDLVNY----FEVGATYYFNKNMSTYV 311
                                                                                                                                                                                  PCTUS4.08327-6
Sequence 6, Application PC/TUS9408327
GENERAL INFORMATION:
APPLICANT: The Rockefeller University
APPLICANT: 1230 YORK Avenue
APPLICANT: New YORK, New YORK 10021
APPLICANT: United States of America APPLICANT: 12103 Indian Creek Court
APPLICANT: 12103 Indian Creek Court
APPLICANT: Beltsville, Maryland 20705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1438.006PC00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/096,182
FILING DATE: 23 July 1993
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W,
REGISTATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 1438.006FCO
TELECOMUNICATION INFORMATION:
TELEPHONE: (202) 371-250
INFORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                                                                                                                                                                                                 United States of America
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : 363 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; MOLECULE TYPE: protein
PCT-US94-08327-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE: D.C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Simi
Matches 83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62
                                                                                                                                                                        RESULT 13
g
                                                  õ
                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

Gaps

us-09-490-291-9.rai

```
APPLICANT: Tai, Joseph Y.
APPLICANT: Tai, Joseph Y.
APPLICANT: Liang, Shur Mei
APPLICANT: Liang, Shur Mei
APPLICANT: Liang, Shur Mei
APPLICANT: Hronowski, Lucjan J.J.
APPLICANT: Pullen, Jeffrey K.
TITLE OF INVENTION: Method for the High Level Expression,
TITLE OF INVENTION: Pourification and Refolding of the Outer Membrane Group B
TITLE OF INVENTION: Porin Proteins from Neisseria meningitidis
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
                                                                                                                               122 LISVRYDSPEFAGLSGSVQYAL-----NDNAGRHNSESYHAGFNYKNGGFFV--QYGG 172
                                                                                                                                                                                                                                                                                                                             :| :| :| :| :| :| :| 25 SQTEVAATLAYRFGGNUTPRVSY----AHGFKGLVDDADIGNEYDQVVVGAEYDFSKRTS 279
84 AFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGGD----TAYSDDFFVGRVGG---- 135
                                            74 SFIGLK-GGFGKLRVGRLNSVLKDT-----GDINPWDSKSDYLGVNKIAEPEAR 121
                                                                                                --VATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISYEYEGFGIVGAYGA 193
                                                                                                                                                                                             194 ADRTNIQEAOPLGNGK-KAEQWATGLKYDANNIYLAANYGETRNATPITNKFTNTSGFAN 252
                                                                                                                                                                                                                                                                                                253 KTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEG-IGDVDLVNYFE---VGATYYFNKNMS 308
                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Ave., NW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: Esmond, Robert W.
REGISTRATION UNDRER: 32,893
REFERENCE, DOCKET NUMBER: 1438.0060001
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,264
FILING DATE: 28-APR-1995
APPLICATION NUMBER: 08/096,182
FILING DATE: 23-JUL-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
US-08-877-109-2
: Sequence 2. Application US/08877109
: Patent No. 5747287
: GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Blake, Milan S. APPLICANT: Tai, Joseph Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : 309 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; TOPOLOGY: linear; MOLECULE TYPE: protein US-08-877-109-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: Washington STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20002
                                                                                                                                                                                                                                                                                                                                                                                              309 TYV 311
                                                                                                                                                                                                                                                                                                                                                                                                                                               280 ALV 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                             136
                                            g
                                                                                        δ
                                                                                                                                    g
                                                                                                                                                                                             ö
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                              ò
```

8.2%; Score 147.5; DB 1; Length 309;

Query Match

```
17;
                                                                                                                       84 AFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGGD----TAYSDDFFVGRVGG---- 135
                                                                                                                                              122 LISVRYDSPEFAGLSGSVQYAL-----NDNAGRHINSESYHAGFNYKNGGFFV--QYGG 172
                                                                                                                                                                                                                                                            ADRINLQEAQPLGNGK-KAEQWATGLKYDANNIYLAANYGETRNATPITNKFTNTSGFAN 252
                                                                                                                                                                                                                                                                                --VATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISYEYEGFGIVGAYGA 193
                                                                                                                                                                                                                                                                                                                               253 KTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEG-IGDVDLVNYFE---VGATYYFNKNMS 308
                                                    NGDMTY------ARLGFKGETQINSDLTGYGQWEYNFQGNNSEGADAQTGNKTRL 83
                                                                                    19 NGQVTEVTTATGIVDLGSKIGFKGQEDLGNGLKAIWQVE---QKASIAGTDSGWGN--RQ 73
                     Indels
   5.7e-06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Search completed: December 18, 2002, 16:23:14 Job time : 8.85111 secs
   Pred. No.
Best Local Similarity 25.1%; Pr
Matches 76; Conservative 46;
                                                                                                                                                                                                                                                                                                                                                                                                    309 TYV 311
                                                                                                                                                                                                                                                                                                                                                                                                                                      280 ALV 282
                                                    35
                                                                                                                                                                                        136
                                                                                 셤
                                                                                                                       ò
                                                                                                                                                       셤
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                             a
```

This Page Blank (uspto)

```
(without alignments)
1213.493 Million cell updates/sec
                                                                                                                                December 18, 2002, 16:21:10 ; Search time 4.71067 Seconds
                                                                                                                                                                                                                     US-09-490-291-9
1809
1 AEIXNKDGNKVDLYGKAVGL.....NKLGVGSDDTVAVGIVYQFA 341
                                                                                                                                                                                                                                                                                                                                                                                                                              106657
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                  106657 seqs, 16763532 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                      OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                               Perfect score:
                                                                                                                                                                                                                                                                                                                   Scoring table:
                                                                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                                                  Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Database
                                                                                                                                   Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:* /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:* /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

/cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/USOB_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/USOB_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO7_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/USO7_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/PUB.PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/USOB_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/USOB_PUBCOMB.pep:*

Description	Sequence 24, Appl Sequence 381, Appl Sequence 70, Appl Sequence 1, Appl Sequence 192, Appl Sequence 118, Appl Sequence 12, Appl Sequence 22, Appl Sequence 8, Appl Sequence 98, Appl Sequence 14, Appl Sequence 146, Appl Sequence 146, Appl Sequence 111, Appl
SUMMARIES	US-09-905-176-24 US-09-912-020-381 US-09-912-020-340 US-08-813-666A-2 US-08-811-132-192 US-09-811-132-192 US-09-811-132-192 US-08-813-666A-2 US-08-834-666A-2 US-08-834-666A-2 US-09-813-666A-4 US-09-915-242-5131 US-09-915-242-5131 US-09-924-097-14 US-09-924-097-14 US-09-924-66A-131 US-09-924-097-14 US-09-924-097-14
DB	100 100 100 100 100 100 100 100 100
% Query Match Length DB	222 2420 2420 2420 2520 2520 2520 2520 2
å Query Match	1,000,000,000,000,000,000,000,000,000,0
Score	928. 1113. 112. 110. 110. 110. 110. 100. 100. 100
Result No.	100 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Sequence 10, Appl Sequence 166, App	Sequence 2, Appli	Sequence 180, App	Sequence 6, Appli	Sequence 15, Appl	Sequence 19, Appl	Sequence 138, App	Sequence 5106, Ap	Sequence 26, Appl	Sequence 86, Appl	Sequence 364, App	Sequence 9, Appli	Sequence 174, App	Sequence 4, Appli	Sequence 3, Appl1	Sequence 12, Appl	Sequence 23, Appl	Sequence 8, Appli	Sequence 2, Appl1	Sequence 362, App	Sequence 33, Appl	Sequence 10, Appl	Sequence 138, App	Sequence 299, App
٦	9 US-10-092-880-2 10 US-09-815-242-10384		_		_				_		0	10 US-09-881-752A-174		10 US-09-350-259-3	9 US-09-895-913A-12	0	10 US-09-119-900-8	_			9 US-10-092-880-10	9 US-09-895-913A-138	9 US-10-108-605-299
672 595	1536	700	1430	272	934	1290	815	410	715	878	2122	470	1152	1153	1974	839	271	518	793	2353	1600	285	265
5.1		5.1	5.0	5.0	5.0	5.0	4.9		4.9	4.9	4.9	4.9	4.9	4.9			4.8	4.8	4.8	4.8	4.8	4.7	4.7
93	92.5	91,5	90.5	90	90	06	89.5	88	89	89	89	88.5	88	88	88	87.5	. 87	87	87	87	98	85.5	82
20	22	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

		1 US/09905176	541		INSTITUTE OF TECHNOLOGY	r A.	TITLE OF INVENTION: METHOD FOR DETERMINING THREE-DIMENSIONAL PROTEIN STRUCTURE FRO	ROTEIN SEQUENCE		IBER: US/09/905,176	1002-04-05	rR: US 60/218,016	1-07-12	9?	11on 3.1		
KESOLT 1	US-09-905-176-24	; Sequence 24, Application US/09905176	; Patent No. US20020150906A1	; GENERAL INFORMATION:	; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY	; APPLICANT: Debe, Derek A.	; TITLE OF INVENTION: METHOD FOR	; TITLE OF INVENTION: PROTEIN SEQUENCE	; FILE REFERENCE: 265/297	; CURRENT APPLICATION NUMBER: US/09/905,176	; CURRENT FILING DATE: 2002-04-05	; PRIOR APPLICATION NUMBER: US 60/218,016	; PRIOR FILING DATE: 2000-07-12	; NUMBER OF SEQ ID NOS: 26	; SOFTWARE: Patentin version 3.1	; SEQ ID NO 24	: LENGTH: 322

ALIGNMENTS

Indels 27; Gaps DB 10; Length 322; Query Match 51.2%; Score 925.5; DB 10; Best Local Similarity 57.3%; Pred. No. 1.1e-76; Matches 192; Conservative 35; Mismatches 81;

) ORGANISM: Klebsiella pneumoniae US-09-905-176-24

.; œ

70 SEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGGDTAYSDDFF 129 130 VGRVGGVATYRNSNFFGLVDGLNFAVQYLGKN-----ERDT----ARRSNGDGVGGSIS 179 10 KVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQWEYNFQGNN 69 ö 음 à g à 셤

180 YE-YEGFGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNAT 238 ò g

```
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18;
                                                                                                                                                                                                                                                                                                              APPLICANT: Zyskind, Judith
APPLICANT: Zyskind, Mari L.
APPLICANT: Ohlsen, Kari L.
APPLICANT: Tradick, John
APPLICANT: Tradick, John
APPLICANT: Froslich, Jamea
APPLICANT: Froslich, Jamea
APPLICANT: Froslich, Jamea
APPLICANT: Tamamoto, Robert T.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
TITLE OF INVENTION: GOVERNOR BREEN OF 101-07-23
FRIOR FILING DATE: 2001-07-23
PRIOR FILING DATE: 1999-01-27
PRIOR FILING DATE: 1999-01-27
PRIOR FILING DATE: 1999-01-27
NUMBER OF SEQ ID NOS: 485
SOFTWARE: FastSEQ for Windows Version 3.0
                     233 R----AGSLGFANKAQNFEVAAQYQFDFGLRPSVAYLQSKGKDLNGYGDQDILKYVDVG 287
239 PITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVG 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          213 YF----GGNFNGDNGTINY------DDDVNGIMVGV---DTKIDGNNAKWIVGAAA--- 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79 NKTRLAFAGLKYADVGSFDYGRNYGVYYDALGYTDMLPEFGGDTAYSDDFFVGRVGGVAT 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               256 ------IYSSAH 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             139 YRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISYEYEGFGI------ 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          188 --VGAYGAADRT-NLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPITNKF 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     338 GYVTPYGSVSGLFQSGDDYQLSNDMKVD----GQSYDSMRYELGVDAGYT------F 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                T-----YSEDQALTPYFKLAYVYDDSNNDNDVNGDSIDNGTEGSAVRV 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22 YFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQWEYNFQGNNSE---GADAQTG 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   245 TNTSGFANKTQDVLLVAQYQFDFGLRP--SIAYTKSKAKDVEGIGDVDLVNYFE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6.2%; Score 113; DB 10; Length 467;
22.8%; Pred. No. 0.0079;
tive 31; Mismatches 114; Indels 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLGTQFSFTKNFSAYTD-----ANYLGGGDVDQDWSANVGVKY 465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            297 -VGATYYFNKNMSTYVDYIINQIDSDNKLGVGSDD---TVAVGIVY 338
                                                                                                                 299 ATYYFNKNMSTYVDYIINQIDSDN---KLGVGSDD 330
                                                                                                                                                                                                                                                      Sequence 381, Application US/09912020 Patent No. US20020045592A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 70, Application US/10029180 Publication No. US20020182708A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 22.8%
Matches 79; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; TYPE: PRT
; ORGANISM: E. COli
US-09-912-020-381
                                                                                                                                                                                                                                 US-09-912-020-381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-10-029-180-70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          428
                                           g
                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
ò
                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
```

```
APPLICANT: Holtzman, Doug
APPLICANT: Holtzman, Doug
APPLICANT: Holtzman, Doug
APPLICANT: Mina, G. Todd
APPLICANT: Sherman, Anir
APPLICANT: Sherman, Anir
APPLICANT: Slava, Jeffry
APPLICANT: Sherman, Anir
APPLICANT: Sherman, Anir
APPLICANT: Lahang, Lixin
APPLICANT: Lahang, Lixin
FILE REFERENCE: MCC-004
CURRENT APPLICATION NUMBER: US/10/029,180
CURRENT APPLICATION NUMBER: US 60/257,431
PRIOR PILING DATE: 2000-12-22
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 138
SOFTMARE: FastSEQ for Windows Version 4.0
SEQ ID NO 70
LENTH: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 340, Application US/09912020

Patent No. US20020045592A1

GENERAL INFORMATION:

APPLICANT: Zyskind, Judith

APPLICANT: Trawick, John

APPLICANT: Trawick, John

APPLICANT: Froelich, Jamie M.

APPLICANT: Trawick, John

APPLICANT: Trawick, John

APPLICANT: Trawick, John

APPLICANT: Trawick, John

APPLICANT: Treelich, Jamie M.

APPLICANT: Treelich, Jamie M.

APPLICANT: Treelich, Jamie M.

APPLICANT: Wu, H. Howard

TITLE OF INVENTION: ESCHERICHIA COLI

TITLE OF INVENTION: ESCHERICHIA COLI

TITLE OF INVENTION: ESCHERICHIA COLI

TITLE OF INVENTION: UNMERR: US/09/912,020

CURRENT FILING DATE: 2001-07-23

PRIOR FILING DATE: 2000-01-27

PRIOR FILING DATE: 1999-01-27

PRIOR FILING DATE: 1999-01-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 GQQGGSSDYYAGQQHQQGGYGQQQGGSSDY-YAGQQHQQHGYCQHDQNRQGGYEQQQHGA 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  112 TOMLPEFGGDTAYSDDFFVGRVGGVATYR-NSNFFGLVDG--LNFAVQYLGKNERDTARR 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 NFQGNNS------EGADAQTGNKTRLAFAGLKYAD--VGSFDYGRNYGVVYDALGY 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 YNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQWEY 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6.2%; Score 112; DB 9;
24.8%; Pred. No. 0.0035;
tive 24; Mismatches 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQ ID NOS: 485
SOFWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 340
LENGTH: 1426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          169 S-----NGDGVGGSISYEYEGF 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              179 SPPQYGGPPPSNSGSGGSMMDQLGGF 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OTHER INFORMATION: fungal gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 6.2%
Best Local Similarity 24.8%
Matches 51; Conservative
Cali, Brian M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT
ORGANISM: E. Coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-912-020-340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-10-029-180-70
```

20;

```
357 AEILNQAEQVVKNFEKIPTAFVSDSLGVCYEVQG-GERRGTNPGQVTSNTWGAGCAYVKQ 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  186 --GIVGAYG-AADR----TNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNAT 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       47 GETQINSDLTGYGQMEYNFQGNNSEG-ADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVV 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  106 YDALGYTDMLPEFGG-----DTAYSDDFFVGRVGGVAT--YRNSNFFGLVDGLNFAVQYL 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               462 TTALSKVPNAQSLQNVVSKKNNPYSPQ-----GIETNYYLNQNSYNQIQTIN---QEL 511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             159 GKNERD-----TARRSNG--DGVGGSISYE------YEGF-------185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       239 PITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDL----- 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    572 SASDVWTYGFGADALYNFINDKATNFLGKNNKLSVGLFGGIALAGTSWLNSEYVNLATVN 631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.7%; Score 104; DB 10; Length 848;
22.6%; Pred. No. 0.11;
tive 39; Mismatches 103; Indels 104; Gaps
                                                                                                                                                                                                                                                                                                                                                                                               1 AEIYNKDGNKVDLYGK------AVGLHYFSKGNGENSYGGNGDMTYARLGF-----K 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---- GQWEYNFQGNNSEGADA 75
                                                                                                                                                                                                                                                                                                     Ouery Match 5.8%; Score 104.5; DB 8; Length 708; Best Local Similarity 20.8%; Pred. No. 0.079; Matches 80; Conservative 53; Mismatches 145; Indels 107;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Bhatia, Ajay
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Probst, Peter
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 21011.466908
CURRENT APPLICATION NUMBER: US/09/841,132
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSEQ for Windows Version 3.0/4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29 ENSYGGNGDMTYARLGFKGETQINSDLTGY-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  681 NTNYYSFMCAELKYRRLYSVYLNYV 705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 192, Application US/09841132 Patent No. US20020061848A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        292 -VNYFE-VGATYYFNKNMSTYVDYI 314
                                                                                                                                                                                  NAME/KEY: Signal Sequence LOCATION: 1...19
OTHER INFORMATION:
                            : 708 amino acids amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 5.7% Best Local Similarity 22.6% Matches 72; Conservative
    SEQUENCE CHARACTERISTICS
                                                                                          TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
                                                                            single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; TYPE: PRT
; ORGANISM: Chlamydla
US-09-841-132-192
                          LENGTH: 708 am:
TYPE: amino aci
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-841-132-192
                                                                                                                                                                                                                                                       US-08-834-666A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQ ID NO 192
                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                               19;
                                                                                                                                                                                                                                                                                                                                                    185 FGIVGAYGAADRTNLQEAQPLGNGKKAEQ-----WATGLKYDANNIYLAANYG---- 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ...-ETRNATPITNKFTNTSGFANK-TQDVLLVAQY----- 263
                                                                                                                                                                                                                                                                        529 TGSTROMTWSRY---GQLLAFTDCSGYQTRYEYDRFGOMTAVHREEGISLYRRYDNRGRL 585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  753 RLTGECQTVENPETGELLWQHETKHA-----YNEQGLANRVTPDSLPPVEWLTYGSGY 805
                                                                                                                                                                                  469 KFYYNDGNQLTAVVSPDGLESRREYDEPGRLVSETSRSGETVRYRYDDAHSELPATTTDA 528
                                                                                                                                                                                                                                                                                                                        SEGADAQTGNKTRLAFAGLKYADVGSFDY----GRNYGVVYDALGYTDMLPEFGGDTAY 124
                                                                                                                                                                                                                                                                                                                                                                                                                  SDDFFVGRVGCVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISYEYEG 184
                                                                                                 Gaps
                                                                                                                                                                                                                             32 YGGNGDMTYARLGFKGETQINSDLTGY------GQW-----EYNFQGNN 69
                                             6.1%; Score 110; DB 10; Length 1426;
21.8%; Pred. No. 0.06;
tive 46; Mismatches 143; Indels 126;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Kleanthous, Harold
APPLICANT: Lissolo, Ling
APPLICANT: Tomb, Jean-Francois
APPLICANT: Tomb, Jean-Francois
APPLICANT: Miller, Charles
APPLICANT: Al-Garawi, Amal
TITLE OF INVENTION: 76 kDa Helicobacter Polypeptides and
TITLE OF INVENTION: Corresponding Polynucleotide Molecules
CORRESPONDENCE: 64
CORRESPONDENCE ADDRESS:
                                                                                                                                       2 EIYNKDGNKVDLYGKAVGL----HYFSKG-----NGE------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        806 LAGMKLGGTPLVEYTRDRLHRETVRSFGSMAGSNAAYELTSTY 848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             264 -- QFDFGLRPSIAYTKSK -- AKDVEGIGDVDLVN - YFEVGATY 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OPERATING SISTEM: DOS
SOFTWARE: FASTEED for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/08/834,666A
FILING DATE: 01-APR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: CLARY, PAUL T.
REGISTRATION NUMBER: 30,175
REGISTRATION NUMBER: 30,175
REFRENCE/DOCKET NUMBER: 06132/038001
TELEPHONE: 617-428-0200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/08834666A Patent No. US20020044949A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Clark & Elbing LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE: Clark & Elbing
STREET: 176 Federal Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER REALABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
                                          Query Match
Best Local Similarity 21.8<sup>3</sup>
Matches 88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02110
US-09-912-020-340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CITY: BOST
STATE: MA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-834-666A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
```

20;

```
169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1323 ---YTGF-LAGSWFFKGQYSLGETQ---NDMKTRYGVLGESSASWTSRGVLADALVEYRS 1375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1376 LVGPVRPTFYALHFNPYVEVSYASMK-----FPGFTEOGREARSFEDASL-TNITIPLGM 1429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              600 GVDIQLMEDFVLG-VSGAA-----FLGKMDSQKF------DAEVSRK----GVVGSV 640
493 EDTYGHMGDWSEAKIQ-DGTLVINWNPTGYRLDPQKAGALVFNALWE------EGAVL 543
                                                    544 SALKNAR--FAHNLTAQRMEFDYSTNVWGFAFG--GFRTLSAENLVAIDGYKGAYGGASA 599
                                                                                        121 --DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSI 178
                                                                                                                                                                          -----NNIYLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGL 269
                              76 QTGNKTRLAFAGLKYADVGSFDYGRN-YGVVYDALGYTDMLPE------FGG--- 120
                                                                                                                                                     179 SYEYEGFGIVGAYGAADRTNLQEAQPLGN-----GKKAEQWAT-GLKYDA----- 222
                                                                                                                                                                                                                                              694 LVGPVRPTFYALHFNPYVEVSYASMK ----FPGFTEQGREARSFEDASL-TNITIPLGM 747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         76 QTGNKTRLAFAGLKYADVGSFDYGRN-YGVVYDALGYTDMLPE------FGG--- 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 --DIAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSI 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----NNIYLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGL 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 5.7%; Score 104; DB 10; Length 1530;
Best Local Similarity 22.6%; Pred. No. 0.23;
Matches 72; Conservative 39; Mismatches 103; Indels 104; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29 ENSYGGNGDMTYARLGFKGETQINSDLTGY------GQWEYNFQGNNSEGADA 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SYEYEGFGIVGAYGAADRTNLOEAQPLGN------GKKAEQWAT-GLKYDA----
                                                                                                                                                                                                                                                                                                                                                                                : : |: | : | : | 1430 KFELAFIKGQFSEVNSLG 1447
                                                                                                                                                                                                                                                                                               : :|: | : :| :|
748 KFELAFIKGQFSEVNSLG 765
                                                                                                                                                                                                                                                                             270 RPSIAYTKSKAKDVEGIG 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RPSIAYTKSKAKDVEGIG 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
ORGANISM: Chlamydia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-841-132-178
                                                                                                                                                                                                                                                                                                                                                                       US-09-841-132-178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       O ID NO 178
LENGTH: 1530
                                                                                                                      셤
                                                                                                                                                                                g
                                                                                                                                                                                                                                              g
                                ò
                                                                                        ò
                                                                                                                                                      ð
                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 NFQGNNS--EGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGGD 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        506 DFONNGAMNGIGVOVGYK---OFFGKK------RNWGLRY--YGFFDY----N 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122 TAYSDDFFVGRVGGVATY------RNSNFFGLVDGLNFAVQYLGKNERDTARR 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   226 YLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEG 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SNGDGVGGSISYEYEGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYD---ANNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----SWINSQOVNLTMMNGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 122;
                                                                                                              APPLICANT: Lissolo, Ling
APPLICANT: Tomb, Jean-Francois
APPLICANT: Miller, Charles
APPLICANT: Al-Garawi, Amal
TITLE OF INVENTION: 76 kDa Helicobacter Polypeptides and
TITLE OF INVENTION: Corresponding Polynucleotide Molecules
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 691;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/08/834,666A
FILING DATE: 01-APR-1997
CLASSIFICATION: 424
ATTORNEY_AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,175
REFERENCE/POCKET NUMBER: 06132/038001
TELEPHONE: 617-428-0200
TELEPHONE: 617-428-0200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 5.7%; Score 103; DB 8; Best Local Similarity 21.6%; Pred. No. 0.1; Matches 60; Conservative 27; Mismatches 69.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          286 IGDVDL------VNYFE-VGATYYFNKNMSTYVDYI 314
US-08-834-666A-12; Sequence 12, Application US/08834666A; Patent No. US20020044949A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                          STREET: 176 Federal Street CITY: Max
                                                                                        APPLICANT: Kleanthous, Harold APPLICANT: Lissolo, Ling
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----GLFG-----GFALAGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KET: Signal Sequence
LOCATION: 1...18
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                    ZIF: V. COMPUTER FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM COMPALIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: 1 SEQUENCE CHARACTERISTICS: LENGTH: 691 amino acids TYPE: amino acid STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM COM
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-834-666A-12
                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

셤

```
286 IGDVDL-----VNYFE-VGATYYFNKNMSTYVDYI 314
                                                                                                                                                          Sequence 8, Application US/08834666A Patent No. US20020044949A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                 E: Clark & Elbing LLP
176 Federal Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; NAME/KEY: Signal Sequence
; LOCATION: 1...20
; OTHER INFORMATION:
US-08-834-666A-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 711 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEFAX: 617-428-7045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                        STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE:
                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----RNSNFFGLVDGLNFAVQYLGKNERDTARR 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SNGDGVGGSISYEYEGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYD---ANNI 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      YLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEG 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             64 NFQGNNS -- EGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGGD 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----SWLNSQQVNLTMMNGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          69; Indels 122;
                                                                                                                                                                       APPLICANT: LISCALO, Ling
APPLICANT: LISCALO, Ling
APPLICANT: Tomb, Jean-Francois
APPLICANT: Tomb, Jean-Francois
APPLICANT: Miler, Charles
APPLICANT: Miler, Charles
APPLICANT: Miler, Charles
TITLE OF INVENTION: 76 kDa Helicobacter Polypeptides and
TITLE OF INVENTION: Corresponding Polynucleotide Molecules
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Blbing LLP
STREET: 176 Federal Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5.7%; Score 103; DB 8; Length 691; 21.6%; Pred. No. 0.1; tive 27; Mismatches 69; Indels 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/08/834,666A
FILING DATE: 01 APR-1997
651 QHGIELGFKIPTINTNYSFMGAKLEYRRMYSLFLNYV 688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,175
REFERENCE/DOCKET NUMBER: 06132/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEPHONE: 617-428-7045
                                                                                            Sequence 22, Application US/08834666A Patent No. US20020044949A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----GLFG-----GFALAGT----
                                                                                                                                                          Kleanthous, Harold
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122 TAYSDDFFVGRVGGVATY-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO: 22:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: Signal Sequence LOCATION: 1...18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 691 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          60; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein
FRAGMENT TYPE: internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCATION: 1...18
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                            Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                    02110
                                                                            US-08-834-666A-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-834-666A-22
                                                                                                                                                                                                                                                                                                                                                                                               STATE: MA
                                                                                                                                                        APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local
Matches 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             617
                                                         RESULT 9
```

```
223 NNIYLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKD 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                132 RVGGVATYRNSNFFGLVDGLNFAV---QYLGKNERDTARRSNGDGVGGSISYEYEGFGIV 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    189 GAYGAADRINLQEAQPLGNGKKAEQWATGLKYD-------A 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     561 -YYGFVDYNHTYNKSQFFN-SDSDVWTYGVGSDLLVNFINDKATKHNKISFGAFGGIQLA 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         74;
                                                                                                                                                                                                                                              APPLICANT: Kleanthous, Harold
APPLICANT: Lissolo, Ling
APPLICANT: Lissolo, Ling
APPLICANT: Tomb, Jean-Francois
APPLICANT: Miler, Charles
APPLICANT: Allery Charles
APPLICANT: Al-Garwah, Amal
TITLE OF INVENTION: 76 kDa Helicobacter Polypeptides and
TITLE OF INVENTION: Corresponding Polynucleotide Molecules
WUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark 6 Elbing LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 711;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      75; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ZIP: COMPUTER READBLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: Diskette
COMPUTER: Diskette
COMPUTER: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSED for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,666A
FILING DATE: 01.APRR.1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30.175
REFERENCE/DOCKET NUMBER: 06132/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617.428-2200
Query Match 5.6%; Score 102; DB 8
Best Local Similarity 20.8%; Pred. No. 0.13;
Matches 46; Conservative 26; Mismatches
```

.. 6

à g

g ò q ò

```
; ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 01 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: Boston STATE: MA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-834-666A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : | | :: | | :: | 1.084 -----ETFTCQMGSTVLFKEG------ALTVNKGGISQGELTGGGNLNVTGGTLAIE 1129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1130 ----GLNARYNALTSISPNAEVSLDNTQGLGRGNIANDGLLTLK----NV-----TGELR 1176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1177 NS--ISGK-----GIVSATARTDVELDGDNSRFVGQFNIDTGSALSV-----NEOKNL 1222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              122 TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISYE 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182 YEGFGIVGAYGAAD-----RTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETR 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63 -YNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGGD 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NATPITNKFTNTSGFANKT--QDVLL-----VAQYQFDFGLRPSIAYTKSKAKDVEGI 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14 YG-KAVGLHYFSK---GNGENSYGGNGDMTYARLGFKGETQIN----SDLTGYGQWE-- 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1045 TYVQMGTLRTDADGALGNTRELNISN-------AAIVDLNGSTQTV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.6%; Score 101; DB 10; Length 1325; 21.4%; Pred. No. 0.36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       287 GDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKLGVG----SDDTVA 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 98, Application US/10007693
Patent No. US20020146776a1
GENERAL INFORMATION:
APPLICANT: BALAIA, Ajay
APPLICANT: BALAIA, Ajay
APPLICANT: Probst, Peter
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 210121.515C2
CURRENT FILING DATE: 2001-12-05
NUMBER OF SEQ 1D NOS: 157
                                                                                                                                      Sequence 304, Application US/09741669
Patent No. US20020022718A1
GENERAL INFORMATION:
GENERAL INFORMATION:
HAPPLICANT: Oblsen, Kari L.
HAPPLICANT: Oblsen, Kari L.
HAPLICANT: Zyskind, Judith W.
TITLE OF INVENTION: Genes identified as required for TITLE OF INVENTION: PRO11feration of E. coli
FILE REFERENCE: ELITRA.009A
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/173005
PRIOR PILING DATE: 1999-12-23
                            668 HSAQHGMELGVKIPTINTNYYSLLGTTLQYRRLYSVYLNYV 708
  -----VNYFE-VGATYYFNKNMSTYVDYI 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             45; Mismatches 133;
                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQ ID NOS: 481
SOFWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 304
LENGTH: 1325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; ORGANISM: Escherichia coli
US-09-741-669-304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 75; Conservat
283 VEGIGDVDL--
                                                                                                                       US-09-741-669-304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 12
US-10-007-693-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 1531
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQ ID NO 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       236
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

```
20;
                                                                                                                                                                                                                                 1167 IHVVTPEIEEDTYGHMGDWSEAKIQ-DGTLVISWNPTGYRLDPQKAGALVFNALWE---- 1221
                                                                                                                                                                                                                                                                                                   118 ---FGG----DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRS 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                220 YDA-------NNIYLAANYGETRNATPITNKFTNTSGFANKTQDVLLV 260
                                                     Indels 104; Gaps
                                                                                                  20 LHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGY------GQWEYNFQ 66
                                                                                                                                                                                                    67 GNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRN-YGVVYDALGYTDMLPE-----
                                                                                                                                                                                                                                                                                                                                                                                                        170 NGDGVGGSISYEYEGFGIVGAYGAADRTNLQEAQPLGN------GKKAEQWAT-GLK
     Length 1531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 4, Application US/08034666A
Petent No. US2002004499A1
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Lissolo, Ling
APPLICANT: Tomb, Jean-Francois
APPLICANT: Miller, Charles
APPLICANT: Al-Garawi, Amal
TITLE OF INVENTION: 76 kDa Helicobacter Polypeptides and
TITLE OF INVENTION: Corresponding Polynucleotide Molecules
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
Query Match 5.5%; Score 100; DB 12; Best Local Similarity 22.0%; Pred. No. 0.53; Matches 72; Conservative 42; Mismatches 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OFFIGURE SYSTEM: DOS SOFTWARE: FastSEQ for Windows Version 2.0 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/834,666A FILING DATE: 01-APR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,175
REFERENCE/DOCKET NUMBER: 06132/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFRAX: 617-428-7045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1422 TNITIPLGMKFELAFIKGQFSEVNSLG 1448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    261 AQYQFDFGLRPSIAYTKSKAKDVEGIG 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Clark & Elbing LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET: 176 Federal Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 733 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKEtte
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
```

```
Query Match
56
                                   g
                                                                                                     g
                                                                                                                                         ò
                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                      임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                   à
                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 146, Application US/09793306

Fatent No. US20020098200A1

GENERAL INFORMATION:
APPLICANT: Skeiky, Yasir
APPLICANT: Owendale, Pamela
APPLICANT: Overdale, Pamela
APPLICANT: Overdale, Pamela
APPLICANT: Overdale, Pamela
APPLICANT: Corriva Corporation
TITLE OF INVENTION: of Tuberculosis
TITLE OF INVENTION: of Tuberculosis
TITLE OF INVENTION: of Tuberculosis
CURRENT PELING DATE: 2001-02-26
FRIOR APPLICATION NUMBER: US 60/185,037
FRIOR FILING DATE: 2000-02-25
FRIOR APPLICATION NUMBER: US 60/23,828
FRIOR APPLICATION NUMBER: US 60/223,828
FRIOR APPLICATION DATE: US 60/223,828
                                                                                                                                                                                                                  17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14;
                                                                                                                                                                                                                                                                                                                        ALGYTDMLPEFGG-----DTAYSDDFFVGRVGGVAT--YRNSNFFGLVDGLNFAVQYLGK 160
                                                                                                                                                                                                                                                                                                                                                489 ALSKVPNAQSLQNVVSKKNNPYSPQ-----GIETNYKLNQNSYNQIQTIN---QELGR 538
                                                                                                                                                                                                                                                                                                                                                                                                                    186 GIVGAYG-AADR----TNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI 240
                                                                                                                                                                                                                                                                           599 SDVWTYGFGADALYNFINDKATNFLGKNNKLSLGLFGGIALAGTSWLNSEYVNLATVNNV 658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         241 TNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDL-----V 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | | : | : | : | : | 59 YNAKMNVANF-------QFLFNMGVRMNLARSKKKGSDHAAQHGIELGLKIPTINT 707
                                                                                                                                                                                                                                                  49 TQINSDLIGYGQWEYNFQGNNSEG-ADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYD 107
                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; OTHER INFORMATION: Description of Artificial Sequence:mTTC#3-His
US-09-793-306-146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91;
                                                                                                                                                                                                                  92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 597;
                                                                                                                                                                             5.5%; Score 99; DB 8; Length 733; ilarity 20.4%; Pred. No. 0.26; Conservative 44; Mismatches 121; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 5.4%; Score 98.5; DB 10;
Best Local Similarity 21.6%; Pred. No. 0.22;
Matches 59; Conservative 30; Mismatches 93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |||:::||
|NYYSFMGAELKYRRLYSVYLNYV 730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NYFE-VGATYYFNKNMSTYVDYI 314
                                                                                  NAME/KEY: Signal Sequence LOCATION: 1...20
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQ ID NOS: 164
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 146
LENGTH: 597
               TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
STRANDEDNESS: single
                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
US-09-793-306-146
                                                                                                                                         US-08-834-666A-4
                                                                                                                                                                               Query Match
                                                                                                                                                                                                 Local
                                                                                                                                                                                               Best Loca
Matches
                                                                                                                                                                                                                                                                                                                        108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    708
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
```

```
19;
                                                                                                                                                                                                                                                                                                                             134 GGVAT-YRNSNFF-----GLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 EYEGFGIV-----TGLK--GAYGAADRTNLQEAQPLGNGKKAEQWA----TGLK-- 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   457 GFANAGSINTGFANSGSINTGGFDSGDRNT-----GFGSSVDQSVSSSGFGNTGMNSS 509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96 FDYGRNYGVVYDALGYTD----- 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :| || || : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || : 
                                                                                                                           310 GSSNIGFGNNGDNNLG-LGNNGNNNIGFGLTGDNLVGIGALNSGIGNLGFGNSGNNNIGF 368
                                                                                                                                                                                                                                               74 DAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGGDTAYSDDFFVGRV 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       52 IDKEFSFKLGGRLQAD---YGRFD-GYYTNNGNTADAAYFRRAYLEFGGTAYRDWKYQIN 107
            73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            39 TYARLGFKGETQINSDLTGYGQWEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVG---S 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GNGENSYGGNGDMTYARLGFKGETQINSDLT------GYGQWEYNFQGNNSEGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 10; Length 440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Haselbeck, Kobert
APPLICANT: Chisen, Kari L.
APPLICANT: 2yskind, Judith W.
APPLICANT: 2yskind, Judith W.
APPLICANT: Trawick, John D.
APPLICANT: Trawick, John D.
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA, 011A
CURRENT APPLICANTON NUMBER: U5/09/815,242
CURRENT APPLICATION NUMBER: 60/191,078
PRIOR APPLICATION NUMBER: 60/201,727
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR PILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/207,737
PRIOR APPLICATION NUMBER: 60/207,737
PRIOR PILING DATE: 2000-11-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches 145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.3%; Score 96.5; DB 21.0%; Pred. No. 0.23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              220 --YDANNIYLAANYGETRNATPITNKFTNTSGF 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            : | : | : | | : | | 510 GFFNTGNV--SAGYGNNGDVQSGINN-TNSGGF 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 5131, Application US/09815242 Patent No. US20020061569A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Haselbeck, Robert
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 87; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-815-242-5131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; ORGANISM: Pseu
US-09-815-242-5131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Fast:
SEQ ID NO 5131
LENGTH: 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
```

VGRVGGVA	166 YDIADWVNDNVGTGIQASSVVGGMAFLSGSVFSENNNDTDGDSVKRYNLRGVFAPLHEPG 225		226 NVVHLGLQYAYRDLEDSAVDTRIRPRMGMRGVSTNGGNDAGSNGNRGLFGGSSAVEGLWK 285		286 DDSVWGLEGAWALGAFSAQAEYLRRTVKAERDREDLKASGYYAQLAYTL 334	NATPITNKFTNTSGFANKTODVLLVAOYOFDFGLRPSIA	TGEPRLYKLDGAKFDTIKPENKEIGAWELFYRYDSIKVEDDNIVVDSATREVG 387	288 DVDLVNYFEVGATYYFNKNMSTYVDYIINQIDS-DNKLGVGSDDTVAVGIVYQF 340	DAKGKTH-TLGVN#YANEAVKVSANYVKAKTDKISNANGDDSGDGLVMRLQYVF 440	Search completed: December 18, 2002, 16:39:31
130	166	150	226	182	286	236	335	288	388	cy Cy
ογ	đ	ογ	ą	ογ	q	δ	đ	ογ	QQ	Sear

Search completed: December 18, Job time : 9.71067 secs

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
```

OM protein - protein search, using sw model

December 18, 2002, 16:13:49; Search time 10.0756 Seconds (without alignments) 3253.588 Million cell updates/sec Run on:

US-09-490-291-9 1809

1 AEIYNKDGNKVDLYGKAVGL......NKLGVGSDDTVAVGIVYQFA 341 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

Total number of hits satisfying chosen parameters:

283224 seqs, 96134422 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	outer membrane por	outer membrane pro	-	outer membrane por		outer membrane por	outer membrane pro	membrane	outer membrane por	outer membrane por		outer membrane por		outer membrane por	outer membrane por	outer membrane pro	outer membrane pro	outer membrane pro	outer membrane por		outer membrane pro	outer membrane por	outer membrane pro	outer membrane por	putaive outer memb	probable outer mem	probable outer mem	outer membrane por	outer membrane pro
SUMMARIES	e e	MMECF	D90755	B85619	D90662	B85513	MMECPE	AH0147	S25520	S25525	AF0543	MMBPP2	MMECPC	836606	MMECNC	D64888	H91016	B85861	AE0616	S51104	843159	AE0690	MMEBPC	AI0789	A59139	D90888	E85729	AE0172	S34263	AC0755
	DB	-	~	~	~	7	Н	7	~	~	7	7	Н	~	-	~	7	7	~	N	~	~	-	7	Н	~	N	7	~	7
	Query Match Length	362	362	362	351	351	351	371	351	320	350	365	367	320	365	377	367	367	363	363	363	383	378	378	378	366	366	360	394	394
di	Query	93.8	98.9	98.9	61.5	61.5	61.4	59.8	59.1	58.5	58.5	58.1	58.1	57.7	57.6	57.2	57.1	57.1	57.0	56.9	56.9	56.8	9.95	56.6	55.3	55.2	55.2	54.8	54.0	54.0
	Score	1805	1789	1789	1113	1113	1110	1082	1070	1058	1058	1021	1021	1044	1042	1035	1033	1033	1031.5	1030	1029.5	1027	1023.5	1023.5	1000.5	998.5	998.5	6	976.5	976.5
	Result No.	-	7	m	4	Ŋ	9	7	α0	o	10	11	12	13	14	15	16	17	18	19	50	21	22	23	24	25	56	27	28	29

outer membrane pro	outer membrane por	hypothetical prote	hypothetical prote	partial probable o	ompF-like porin [i	probable outer mem	probable outer mem	outer membrane por	hypothetical prote	probable outer mem	hypothetical prote	partial probable o	outer membrane por	outer membrane pro	membrane
AH0149	S70847	T46993	F90878	D85740	G84971	F90966	F85814	H64960	390966	G85814	G90878	C85740	B64961	A82299	C64900
7	~	~	~	~	~	~	~	~	~	~	~	~	ď	~	7
374	369	315	255	255	382	214	224	171	191	191	123	123	134	320	96
-4	5	و	6	٦	7	80	8	6	4	σ	و	ဖ	4	-	σ
53.1	<b>4</b> Β.	42.	35.	35.	31.	29	29.	26.	23.	22.	20.	20.	19.	13.	12.
096	685	110	649.5	635.5	573.5	539.5	539.5	486	423	415	372.5	372.5	351	236.5	233.5
30	31	32	33	34	32	36	37	38	33	40	41	42	43	44	45

## ALIGNMENTS

	RESULT 1
	MMECF
_	outer membrane porin ompF precursor - Escherichia coli (strain K-12)
	N/Alternate names: outer membrane protein la; outer membrane protein b
	C;Species: Escherichia coli
	C;Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 01-Mar-2002
	C; Accession: A93449; A91301; A90314; A25029; H64832; A03431
	R:Inokuchi, K.; Mutoh, N.; Matsuyama, S.; Mizushima, S.
	Nucleic Acids Res. 10, 6957-6968, 1982
	A;Title: Primary structure of the ompF gene that codes for a major outer membrane pro
	A; Reference number: A93449; MUID: 83090452; PMID: 6294623
	A;Accession: A93449
	A; Molecule type: DNA
	A; Residues: 1-362 <ino></ino>
	A;Cross-references: GB:J01655; GB:M10311; GB:M10312; NID:q147009; PIDN:AAA24244.1; PI
	A; Experimental source: strain K12
	R; Mutoh, N.; Inokuchi, K.; Mizushima, S.
	FEBS Lett. 137, 171-174, 1982
	A. Mit John Smith and A common of the citems and the action of the citems and the citems are citems and citems are citems are citems and citems are citems are citems and citems are citems are citems and citems are citems and citems are citems are citems are citems and citems are citems are citems are citems and citems are citems are citems and citems are cite

A/Title: Amino acid sequence of the signal peptide of OmpF, a major outer membrane pr A/Reference number: A91301; MUID:82139379; PMID:7037455

A; Recension: A91301; MUID:82139379; PMID:7037455
A; Recension: A91301
A; Molecule type: DNA
A; Residues: 1-37 < AUTY>
A; Residues: 1-37 < AUTY>
Biochem. J. 203, 33-43, 1982
A; Title: Primary Structure of major outer-membrane protein I (ompF protein, porin) of A; Recension: A90314; MUID:8256494; PMID:7049161
A; Reference number: A90314; MUID:8256494; PMID:7049161
A; Rocession: A90314; MUID:8256494; PMID:7049161
A; Rocession: A90314
A; Molecule type: protein
A; Rocession: A90314
A; Molecule type: protein
A; Rocession: A90314
A; Reference number: A91809; MUID:86033642; PMID:2997131
A; Rocession: A25029
A; Molecule type: DNA
A; Residues: 33-63 < NOG>
A; Molecule type: DNA
A; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A; Title: The complete genome sequence of Escherichia coli K-12.
A; Reference number: A64720; MUID:97426617; PMID:9278503
A; A; Title: The complete genome sequence of Escherichia coli K-12.
A; Reference number: A64720; MUID:97426617; PMID:9278503
A; A; Tetle: The complete genome sequence of Escherichia coli K-12.
A; Reference number: A64720; MUID:97426617; PMID:9278503
A; A; Arcession: A64720; MUID:97426617; PMID:9278503

A;Status: nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA
A; Residues: 1.362 <BLAT>
A; Residues: 1.362 <BLAT>
A; Residues: 0.00096; NID: 91787156; PIDN: AAC74015.1; PID: 917871
A; Experimental source: strain K-12, substrain MG1655
C; Genetics:

A;Gene: ompF; tolF; cmlB; coa; cry A;Map position: 21 min C;Complex: homotrimer

300

322

```
A;Accession: B85619
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-362 <STO>
A;Cross-references: GB:AE005174; NID:g12514104; PIDN:AAG55414.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain 0157:H7, substrain EDL933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Outer membrane protein 1a (Ia,b,F) [imported] - Escherichia coli (strain O15 C; Species: Escherichia coli (c; Species: Escherichia coli (c; Species: Escherichia coli (c; Species: Escherichia coli (c; Species: Escherichia species: Escherichia species: Escherichia species: B85619
R; Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, 111er, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousls, Nature 409, 529-533, 2001
A; Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23 AEIXNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGFTQINSDLTGYGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EYEGFGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241 TNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT
                                                                                                                                                                                                                                                               YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; DB 2;
1.3e-121;
                                                                                                                                                                                                                                         YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 98.9%; Score 1789; C
Best Local Similarity 99.4%; Pred. No. 1.3e
Matches 338; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Superfamily: outer membrane protein phoE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gene: ompF
       143
                                                                                                                                                                                                                                                                                     323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     203
                                                     181
                                                                                                203
                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          323
                                                                                                                                                                                                                                                                                                                                                        RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4
                                                                                             g
                                                                                                                                                                                       g
       셤
                                                     ð
                                                                                                                                              ð
                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Outer membrane protein la ECs1012 [imported] - Escherichia coli (strain O157:H7, substra C;Species: Escherichia coli (C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001 (C;Accession: D90755 (C;Accession: D90755 (C;Accession: D90755 (C;Augashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res (B, 11-22, 2001 (C;A) ** Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend A;Reference number: A99629; MUID:21156231; PMID:11258796 (C;A) ** Ascession: D90755 (C;A) ** Ascession: D90755 (C;A) ** Ascession: D90755 (C;A) ** Ascession: Data Makinary (C;
                          small hydrophilic mole
                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule fype: DNA
A; Residues: 1-362 <HAX>
A; Cross-references: GB:BA000007; PIDN:BAB34435.1; PID:g13360471; GSPDB:GN00154
A; Experimental source: strain 0157:H7, substrain RIMD 0503952
C; Genetics:
A; Gene: ECs1012
C; Superfamily: outer membrane protein phoE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYIDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                         DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNCDGVGGSISY 180
                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82
                                                              A Description: receptor for bacterlophage T2
C; Superfamily: outer membrane protein phoE
C; Keywords: membrane protein; porin; trimer
F;1-22/Domain: signal sequence fetatus predicted <SIG>
F;23-362/Product: outer membrane porin ompF #status experimental <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EYEGFGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
<POR>
on: forms aqueous channels that faciliate diffusion of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 362;
                                                                                                                                                                                                                                                            0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                               Length
                                                                                                                                                                                                          Score 1805; DB 1; L
Pred. No. 9.3e-123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 1789; DB 2;
Pred. No. 1.3e-121;
0; Mismatches 2;
                                                                                                                                                                                                                        100.0%; Prec. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Superfamily: outer membrane protein phoE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98.98;
                                                                                                                                                                                                            98.66
                                                                                                                                                                                                       Query Match
Best Local Similarity 100.
Matches 340; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 99.4
Matches 338; Conservative
                                                Function: <REC>
                     : Description:
                                                                                                                                                                                                                                                                                                                                                                                                                                          83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 203
                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

ö

Gaps

;

Length 362 Indels 82

240

300

180 202 0157:H

Han, C

D.J.; May K.; Apoda

(Strain 0157:H7, sub

m

```
A; Cross-references: GB: DB3536; NID: 94902908; PIDN: BAA77910.1; PID: 94902976
A; Experimental source: strain K12
A; Note: this is one of the proteins induced when the E. coll cells are grown under ph a compounds, and some other negatively charged solutes
B; Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A; Ittle: The complete genome sequence of Escherichia coll K-12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A) Description: allows passive diffusion of small hydrophilic molecules with molecular A; Note: induced under phosphate limitation C; Superfamily: outer membrane protein phoE C; Superfamily: outer membrane protein; porin F;1-21/Domain: signal sequence #status predicted <SIG>F;22-351/Product: outer membrane porin phoE #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: DNA
A; Residues: 1-351 <BLAT>
A; Cross-references: GB:AE000132; GB:U00096; NID:g2367098; PIDN:AAC73345.1; PID:g17864
A; Experimental source: strain K-12, substrain MG1655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C;Species: Escherichia coli
C;Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 01-Mar-2002
C;Datession: A03432; B64749
C;Accession: A03432; B64749
J. Mol. Biol. 163, 513-532, 1983
A;Title: Complete nucleotide sequence of phoE, the structural gene for the phosphate A;Reference number: A92893; MUID:83189086; PMID:6341601
                                                                        D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                   WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                     WEAEFAGNKAESDTAQ -- QKTRLAFAGLKYKDLGSFDYGRNLGALYDVEAWTDMFPEFGG 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22 AEIXNKDGNKLDVYGKVKAMHYMS-----DNASKDGDQSYIRFGFKGETQINDQLJGYGR 76
                                                                                                                                                                                                                                                                                TPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       outer membrane porin phoE precursor - Escherichia coll (strain K-12)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 351;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Title: The complete genome sequence of Escherichia coll K-12. A;Reference number: A64720; MUID:97426617; PMID:9278503 A;Accession: B64749 A;Status: nucleic acid sequence not shown; translation not shown
                                                                                                                                                                                                                                                                                                                                                                                       298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61.4%; Score 1110; DB 1; 62.1%; Pred. No. 1.1e-72; 11ve 38; Mismatches 76;
                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 61.4%
Best Local Similarity 62.1%
Matches 213; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
A; Residues: 1-351 < OVE>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Gene: phoE
A;Map position: 6 min
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; Complex: homotrimer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 6
                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                        à
                                                                                                                 용
                                                                                                                                                                          ò
                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                     음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Outer membrane pore protein E (E.Ic./NmpAB) [imported] - Escherichia coli (strain 0157:H7 C.Species: Escherichia coli (c.Species: Escherichia coli (c.Species: Escherichia coli (c.Species: Escherichia coli (c.Species: Escherichia coli (strain 0157:H7 C.Species: Escherichia coli 16-Feb-2001 #text_change 14-Sep-2001 (c.Species) B85513 R.Sperna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew liler, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Astritle: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.
                                                   gend
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Accession: B85513
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-351 <STO>
A;Cross-references: GB:AE005174; NID:912513015; PIDN:AAG54566.1; GSPDB:GN00145; UWGP:203
A;Experimental source: strain 0157:H7, substrain EDL933
DNA Res. 8, 11–22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and (A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D90662
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5
                                                                                                                                                                                                                                                                                                                                                                                                                  5
                                                                                                                                                                                                  A;Cross-references: GB:BA000007; PIDN:BAB33691.1; PID:g13359725; GSPDB:GN00154
A;Experimental source: strain 0157:H7, substrain RIMD 0509952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVYYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121 D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YEYEG--FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 AEIXNKDGNKLDVYGKVKAMHYMSDNDSK----DGDQSYIRFGFKGETQINDQLTGYGR 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 AEIXNKDGNKLDVYGKVKAMHYMSDNDSK----DGDQSYIRFGFKGETQINDQLFGYGR 76
                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 AEIYNKDGNKVDLYCKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          238 TPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                          16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16;
                                                                                                                                                                                                                                                                                                                                                       ; Score 1113; DB 2; Length 35; Pred. No. 6.4e-73; 38; Mismatches 76; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61.5%; Score 1113; DB 2; Length 35
62.1%; Pred. No. 6.4e-73;
tive 38; Mismatches 76; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Superfamily: outer membrane protein phoE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C; Superfamily: outer membrane protein phoE
                                                                                                                                                                                                                                                                                                                                                       Query Match 61.5%;
Best Local Similarity 62.1%;
Matches 213; Conservative 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 213; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                            A; Molecule type: DNA
A; Residues: 1-351 <HAY>
                                                                                                                                                                                                                                                                           A; Gene: ECs0268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
```

ŝ

Gaps

ď

ŝ

Gaps

```
A; Nateriance insures. 325250
A; Molecule type: DNA
A; Recession: $25520
A; Molecule type: DNA
A; Realdues: 1-514 CSPS.
A; Realdues: 1-514 CSPS.
A; Realdues: 1-514 CSPS.
A; Note: this is one of the proteins induced when the E. coli cells are grown under ph compounds, and some other negatively charged solutes
C; Genetics:
A; Gene: phoE
C; Complex: homotrimer
C; Function:
A; Description: allows passive diffusion of small hydrophilic molecules with molecular
A; Note: induced under phosphate limitation
C; Superfamily: outer membrane protein; poin
C; Superfamily: outer membrane protein; poin
F; 1-21/Domain: signal sequence #status predicted <SP>F; 22-351/Product: outer membrane porin phoE #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Outer membrane porin phoE precursor - Salmonella typhimurium
NiAlternate names: phosphate-limitation-inducible outer membrane pore protein phoE
C;Specias Salmonella typhimurium
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Aug-1999
C;Accession: $25525; JQ1937
R;Spterings, G; Edlers, R.; van Hith, B.; Hofstra, H.; Tommassen, J.
Submitted to the EMBL Data Library, August 1992
A;Reference number: $25525
A;Reference number: $25525
A;Reference number: $25525
A;Rolecule type: DNA
A;Residues: 1-350 Aspl.
A;Residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Molecule type: DNA
A;Residues: 1-197;199-350 <SP2>
A;Cross-references: EMBL:X68023
A;Vote: neither the full codon nor the amino acid are given for residue 198 in Figure
C;Genetics:
                                                                                                                                                                                                                                                                                                                                              small hydrophilic molecules with molecular
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            phoE gene and development of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  YEYEG--FGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 351;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 59.1%; Score 1070; DB 2; Best Local Similarity 59.2%; Pred. No. 8.1e-70; Matches 203; Conservative 48; Mismatches 76;
            Reference number: S25520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ribarkhill; J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B. deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M. Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : '
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Cross-references: GB: AL590842; PIDN: CAC90043.1; PID: 915979264; GSPDB: GN00175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         outer membrane protein C2 ompC2 [imported] - Yersinia pestis (strain C092) C; Species: Yersinia pestis (5. Species: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001 C; Accession: AH0147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Outer membrane porin phoE precursor - Citrobacter freundii
C;Species: Citrobacter freundii
C;Species: Citrobacter freundii
C;Species: Citrobacter freundii
C;Accession: S2520
R;Spierings, G.; Ockhuysen, C.; Hofstra, H.; Tommassen, J.
submitted to the EMBL Data Library, August 1992
                                                                                                                                                                                                               DT-AYSDDFFVGRVGCVATYRNSNFFGLVDGLNFAVQYLGKN------ERDTARRSN 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |: :|:| || || ||||||::||||||| || ||::: || DSYTRTDNEMTGRITGVAIYRNIDFFGLVDGLKFSLQYQGKNGAEGEINNGRIDISKQ-N 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAA 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NYGETRNATPIT-----NKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV 283
   DSSAQTDNFMTKRASGLATYRNTDFFGVIDGLNLTLQYQGKNENRDVKKQNGDGFGTSLT 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 AEIYNKDGNKLDLYGKVKALHYFS-----DNTKSDGDKSYVRLGFKGVTQITDELSGYGQ 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EGIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180 YEYEG--FGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGEIRNA
                                                                                                                                                                             238 TPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 371;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     73; Indels
                                                                                                                                                                                                                                                                                                      GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                     Score 1082; DB 2;
Pred. No. 1.2e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  48; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A,Gene: ompC2
C,Superfamily: outer membrane protein phoE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       59.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     212; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-371 <KUR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                      298
                                                                                                                                                                                                                                                                                                                                                           309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 8
ద
                                                          å
                                                                                                                  셤
                                                                                                                                                                             ò
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                      å
                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
```

237

S

```
outer membrane porin ompC precursor - Escherichia coli (strain K-12)
C.Species: Escherichia coli
C.Species: Dec.1990 #sequence_revision 31-Dec.1990 #text_change 01-Mar-2002
C.Accession: A20867; A18885; B25029; E64991
R.Mizuno, T.; Chou, M.Y.; Inouye, M.
J. Blol. Chem. 258, 6932-6940, 1983
A;Title: A comparative study on the genes for three porins of the Escherichia coli ou A;Reference number: A20867; MUID:83213433; PMID:6304064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genes of lambdoid
                                                                                                                                                                                                                                                                                                                    Outer membrane porin lc precursor - phage PA2

C;pecides: phage PA2

A:Note: host Escherichia coli

C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 02-Jul-1998

C;Accession: D25647

B:Blasband, A.J.; Marcotte Jr., W.R.; Schnaitman, C.A.

J. Biol. Chem. 261, 12723-12732, 1986

A;Tille: Structure of the lc and nmpC outer membrane porin protein genes of

A;Reference number: A25647; MUID:86304457; PMID:3017988
                                                                                                               61 WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       231 YGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVD 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24 AEIXNKDSNKLDLYGKVNAKHYFSSNDAD-----DGDTTYARLGFKGETQINDQL,TGFGG 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        177 SISYEYEGFGIVGAYGAADRINLQE-----AQPLGNGKKAEQWATGLKYDANNIYLAAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YEYEG--FGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1. AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNER----DTARRSNGDGVGG
                                                                                         TPITNKFTNTSGFANKT@DVLLVAQY@FDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          291 LVNYFEVGATYYFNKNMSTYVDYIINQIDSDN----KLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Superfamily: outer membrane protein phoE
C;Keywords: membrane protein; porin; trimer
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-365/Product: outer membrane porin lc #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                  ; Pred. No. 2e-68;
35; Mismatches 83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                58.1%; Score 1051; 59.8%; Pred. No. 2e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 59.8%
Matches 211; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: DNA
A; Residues: 1-365 < BLA>
C; Genetics:
                                                194
                                                                                         238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            138
      180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 313
                                                                                                                                                                                                                                                                                 7
                                                셤
                                                                                       ò
                                                                                                                            g
                                                                                                                                                                          à
                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      쉱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Outer membrane pore protein E precursor STY0365 [imported] - Salmonella enterica subsp.

C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C;Accession: AF0543
R;Parkhill, J; Dougan, G; James, K.D; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, J.; Moule, S; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Ouall, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A;Hile: Complete genome sequence of a multiple drug resistant Salmonella enterica serov A;Reference number: AB0502; PMID:11677608
A;Accession: AF0543
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-350 <PAR>
                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Cross-references: GB:AL513382; PIDN:CAD08790.1; PID:g16501606; GSPDB:GN00176 C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                         193
                                                                                                                                                                                                                                                                               120
                                                                                                                                                                                                                                                                                                  121 D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                              180 YEYEG--FGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGO 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                          61 WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  238 TPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16;
C;Superfamily: outer membrane protein phoE
C;Keywords: membrane protein; porin; trimer
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-330/Product: outer membrane porin phoE #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 350;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         82; Indels
                                                                                                                                                    82; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             58.5%; Score 1058; DB 2; 58.0%; Pred. No. 5.9e-69; ive 46; Mismatches 82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 58.5%; Score 1058; DB 2; Best Local Similarity 58.0%; Pred. No. 5.9e-69; Matches 199; Conservative 46; Mismatches 82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Gene: STY0365
C;Superfamily: outer membrane protein phoE
                                                                                                                                                    Conservative
                                                                                                                          Best Local Similarity
Matches 199; Conserv
                                                                                                        Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

7;

```
Rincres, A.; Puente, J.L.; Calva, E. submitted to the EMBL Data Library, August 1993
A; Reference number: 836606
A; Accession: 836606
A; Molecule type: DNA
A; Residues: 1-350 < TOR>
A; Cross-references: EMBL:X74595; NID:9397159; PIDN:CAA52672.1; PID:9397160
A; Experimental source: isolate IMSS-1
A; Cross-references: EMBL:X74595; NID:9397159; PIDN:CAA52672.1; PID:9397160
A; Experimental source: isolate IMSS-1
C; Compounds, and some other negatively charged solutes
C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A Description: allows passive diffusion of small hydrophilic molecules with molecular A; Note: induced under phosphate limitation C; Superfamily: outer membrane protein phoE C; Keywords: membrane protein; porin F;1-20/Domain: signal sequence fatatus predicted <SIG>F;1-350/Product: outer membrane porin phoE #status predicted <OPP>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: S66594
A; Molecule type: DNA
A; Residues: 347-365 < MAH>
A; Cross-references: EMBL: X92587; NID: 91051136; PIDN: CAA63325.1; PID: 91051145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVYYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 AEVYNKNGNKLDVYGKVKAMHYMSDYDSK----DGDQSYVRFGFKGKTQINDQLIGGGR 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   180 YEYEG--FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          238 TPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 57.7%; Score 1044; DB 2; Length 350; Best Local Similarity 57.4%; Pred. No. 6e-68; Matches 197; Conservative 47; Mismatches 83; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Gene: phoE
C; Complex: homotrimer
C; Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                        A;Cross-references: GB:K00541; GB:M10314; GB:M14188; NID:g147007; PIDN:AAA24243.1; PID:grMzuno, T.; Chou, M.Y.; Inouye, M. FEBS Lett. 151, 159-164, 1983
A;Title: DNA sequence of the promoter region of the ompC gene and the amino acid sequence A;Reference number: A18885; MUID:83132326; PMID:6297988
                                                                                                                                                                                                                                                                                                                                                                                                    A; Reference number: A64720; MUID:97426617; PMID:9278503
A; Accession: E64991
A; Molecule acid sequence not shown; translation not shown
A; Residues: 1-367 < RLAT>
A; Across-references: GB:AE000310; GB:U00096; NID:92367131; PIDN:AAC75275.1; PID:91788544; A; App position: 47 min
C; Function:
                                                                                                                                                                                                                                                                                   reco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A Description: one of the E. coll major outer membrane proteins that form passive diffus (Superfamily: outer membrane protein phoE C: Superfamily: outer membrane protein; porin; trimer C: Keywords: membrane protein; porin; trimer F:1-21/Domain: signal sequence #status predicted <SIG>F:22-367/Product: outer membrane porin ompC #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                             ပ္ပ
                                                                                                                                                                                                                                                                                                                                       A; Molecule type: DNA
A; Residues: 32-57 <NOG>
R; Balattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A; Title: The complete genome sequence of Escherichia coli K-12.
                                                                                                                                                                              A; Molecule type: DNA
A; Restatues: 1-40 Anii2>
B; Nogami, T.; Mizuno, T.; Mizushima, S.
J. Bacteriol. 164, 797-801, 1985
A; Title: Construction of a series of cmpF-cmpC chimeric genes by in vivo homologous A; Reference number: A91809; MUID:86033642; PMID:2997131
A; Accession: B29029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .;
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S36606
outer membrane porin phoE precursor - Salmonella typhi
C;Species: Salmonella typhi
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDT------A 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   167 RRSNGDGVGGSISYEYEGFGIVGAYGAADRTNLQE-AQPLGNGKKAEQWATGLKYDANNI 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           226 YLAANYGETRNATPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-E 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 367;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Score 1051; DB 1; Length 36'; Pred. No. 2e-68; 45; Mismatches 69; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    58.1%;
58.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 212; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
A; Molecule type: DNA
A; Residues: 1-367 <MIZ>
                                                                                                                                                      A; Accession: A18885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
```

297 307

ŭ,

us-09-490-291-9.rpr

```
Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        334 VGIVYOF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :|:|||
LGLVYQF 377
                                                                                      Matches 209;
       Query Match
Best Local
                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           37.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        . g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; Species: Escherichia coli

C; Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C; Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C; Accession: D64888

R; Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cc

A, Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A; Title: The complete genome sequence of Escherichia coli K-12.

A; Reference number: A64720; MUID:97426617; PMID:9278503

A; Retens: nucleic acid sequence not shown; translation not shown

A; Residues: 1-377 < BLAT>
       ပ္ပ
                                                                                                                                                                                                                                                            A; Molecule type: DNA
A; Residues: 'MNTYRAVTSFFNNSSKKGLT',1-325,'N',327-347,'EGANKSLI' <BLAT>
A; Cross-references: GB: ABD000160, GB: U00096; NID: 91786751; PIDN: AAC73654.1; PID: 91786765;
A; Experimental source: strain K-12, substrain MG1655
C; Comment: In wild-type strains of E. coli K-12, the nmpC open reading frame is interrup
ain CS348, the IS5 element has been deleted and nmpC is expressed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A.Cross-references: GB:AE000234; GB:U00096; NID:91787633; PIDN:AAC74459.1; PID:91787641.
A.Experimental source: strain K-12, substrain MG1655
C.Superfamally: outer membrane protein phose
C.Superfamally: outer membrane protein; trime
C.Keywords: membrane protein; porin; trime
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-377/Product: probable outer membrane porin b1377 #status predicted <OPP>
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNER---DTARRSNGDGVGG 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           177 SISYEYEGFGIVGAYGAADRINLQE-----AQPLGNGKKAEQWATGLKYDANNIYLAAN 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24 AEIYNKDSNKLDLYGKVNAKHYFSSNDAD----DGDTTYARLGFKGETQINDQLTGFGG 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YGETRNATPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        outer membrane porin b1377 precursor - Escherichia coli (strain K-12)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :24-365/Product: outer membrane porin nmpC #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LVNYFEVGATYYFNKNMSTYVDYIINQIDSDN---KLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 365;
                                                                                                                                                                                                                                    translation not shown
                                                                                                          A; Title: The complete genome sequence of Escherichia coli K-12. A; Reference number: A64720; MUID:97426617; PMID:9278503 A; Accession: G64787 A; Accession: G64787 A; Accession: A; Accession: A; Accession: A; Accession: 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              84; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Map position: 12 min
A; Genome: cryptic lambdoid prophage DLF12
C; Superfamily: outer membrane protein phob
C; Keywords: membrane protein; porin; trimer
F:1-23/Domain: signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 57.6%; Score 1042; DB 1; Best Local Similarity 59.5%; Pred. No. 8.8e-68; Matches 210; Conservative 35; Mismatches 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Gene: nmpC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

```
WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                               RSNGDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIY 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13.1 GGADNPAGVDDKDLVKYADIGATYFNKNMSTYVDYKINLLDEDDSFYAANGISTDDIVA 370
                                                                                                                                                                                                                                      DT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNE------RDTAR 167
                                                                                                                                                                                                                                                                                                                                                                                                   LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVE-- 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVA 333
                                           Gaps
                                                                                                  22 AEVYNKDGNKLDLYGKVDGLHYFSDNSAK----DGDQSYARLGFKGETQINDQLTGYGQ 76
                                                                            1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                              38;
                                         Indels
    Length
                                             78;
    DB 2;
Score 1035; DB 2
Pred. No. 2.9e-67
3; Mismatches 7
                                       42;
57.2%;
56.9%;
                                         Conservative
```

Search completed: December 18, 2002, 16:22:23 Job time : 11.0756 secs

This Page Blank (uspto)

```
December 18, 2002, 16:06:04 ; Search time 5.23408 Seconds (without alignments) 2702.183 Million cell updates/sec
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                         OM protein - protein search, using sw model
                                                                                                                                                           Run on:
```

US-09-490-291-9 1809 1 AEIYNKDGNKVDLYGKAVGL.....NKLGVGSDDTVAVGIVYQFA 341 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

112892 seqs, 41476328 residues Searched:

112892 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_40:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES Reg

Description	esc	93	~		033980 serratia ma	Q01605 citrobacter		P30705 salmonella			P07238 bacteriopha	_	Q01606 klebsiella	_					O52503 salmonella	P76335 escherichia		033507 rahnella aq	Q54471 serratia ma	Q56828 xenorhabdus		vibrio ch				_	216	_	_
ID	PF	PHOE_ECOLI	OMPD_SALTY	PHOE_ENTCL	OMPF_SERMA	PHOE_CITFR	PHOE_SALTI	PHOE_SALTY	OMPC_ECOLI	PHOE_KLEPN	PORI_BPPA2	NMPC_ECOLI	PHOE_KLEOX	OMPN_ECOLI	OMPF_SALTI	OMPF_SALTY	OMPC_KLEPN	OMS2_SALTI	OMPC_SALTY	YEDS_ECOLI	OMS1_SALTI	OMPC_RAHAQ		OMPF_XENNE	PORL_BUCAI	OMPU_VIBCH	OMPL_PHOPR	OP28_HAEIN	OP25_HAEIN	OMB_NEIMB	OP26_HAEIN	OP21_HAEIN	OP29_HAEIN
DB	Н	~4	Н	-	-	Ч	-	Н	Н	Н	-	-	-	æ	-	-	П	-	-	Н	-	П	Н	-	-	Н	Н	7	Н	7	Н	-	٦
Length		351	362	350	374	351	350	350	367	351	365	365	349	377	363	363	363	383	378	397	394	342	376	369	382	341	341	353	371	331	360	359	361
Query Match	99.8	61.4	59.6	59.3	59.3	59.1	58.5	58.5	58.1	58.0	57.8	57.6	57.5	57.2	57.0	57.0	56.9	26.8	56.6	55.2	54.0	49.9	49.4	48.9	31.7	13.1	12.1	4.6	8.4	æ .3	7.8	7.7	7.7
Score	$\sim$	1110	1077.5	1072.5	1072	1070	1058	1058	1051	1050	1046	1042	1041	1035	1031.5	1031.5	1030	1027	1023.5	666	976.5	905	894.5	882	573.5	236.5	218	170.5	151.5	150.5	141.5	139.5	139
esult No.	1	7	m	4	S	9	7	8	6	10	11	12	13	14	12	16	17	18	19	20	21	22	23	24	25	56	27	78	53	30	31	32	33

P18195 neisseria g P20149 haemophilus Q48218 haemophilus Q491y5 neisseria m P20148 neisseria m P30688 neisseria m P55127 neisseria m Q48219 haemophilus Q48219 haemophilus P30687 neisseria m P30687 neisseria m
OMB1_NEIGO OP22_HAEIN OP27_HAEIN PRPC_NEIMB OMB2_NEIGO OMB2_NEIGO OMB2_NEIMC OP2A_HAEIN OMB1_NEIMB OMB1_NEIMB
папапапапапа
348 361 1829 348 331 1829 367 363 331 1310
138.5 138.5 135.5 131.3 131.5 131 131 129.5 129.5
88888844444 4886888884444

### ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Blochemistry 37:15663-15670(1998).
-!- FUNCTION: OMFF IS A PORIN THAT PASSIVE DIFFUSION PORES WHICH ALLOW SMALL MOLECULAR WEIGHT HYDROPHILIC MATERIALS ACROSS THE OUTER MEMBRANE. IT IS ALSO RECEPTOR FOR THE BACTERIOPHAGE T2.
-!- SUBUNIT: HOMOTRIMER.
-!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
-!- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF MUTANT ASP-141.
MEDLINE-95024177; PubMed-7524100;
Jeancher D., Schlimer T., Fourel D., Simonet V., Rummel G., Widmer C.,
Rosenbusch J.P., Pattus F., Pages J.M.;
"Structural and functional alterations of a colicin-resistant mutant
of OmpF porin from Escherichia coli.";
proc. Natl. Acad. Sci. U.S.A. 91:10675-10679(1994).
                                                                                            SEQUENCE OF 33-63 FROM N.A.
MEDLINE-88033642; PubMed-299131;
MOGBMIT T., Mizuno T., Mizushima S.;
"Construction of a series of ompF-ompC chimeric genes by in vivo homoLogous recombination in Escherichia coli and characterization of the translational products.";
J. Bacteriol. 164:797-801(1985).
                                                                                                                                                                                                                                                                                                         STRAIN-K12 / EMG2;
MEDLINE-97443975; PubMed-9298646;
Link A.J., Robison K., Church G.M.;
"Comparing the predicted and observed properties of proteins encoded
in the genome of Escherichia coli K-12.";
Electrophoresis 18:1259-1313(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
MEDLINE-99060042; PubMed-9843370;
Phale P.S., Philippsen A., Kiefhaber T., Koebnik R., Phale V.P.,
SCHILMER T., Rosenbusch J.P.;
Stability of trimeric OmpF porin: the contributions of the latching
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-KI2 / W3110;
MEDLINE-98291876; PubMed-9629924;
Molloy M.P., Herbert B.R., Walsh B.J., Tyler M.I., Traini M.,
Sanchez J.-C., Hochstrasser D.F., Williams K.L., Gooley A.A.;
Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis.";
"Primary structure of major outer-membrane protein I (ompF protein, porin) of Escherichia coli B/r."; Blochem. J. 203:33-43(1982).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
MEDLINE-92375189; PubMed=1380671;
Cowan S.W., Schirmer T., Furmenl G., Steiert M., Ghosh R.,
Pauptit R.A., Janssonius J.N., Rosenbusch J.P.;
"Crystal structures explain functional properties of two E. coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; J01655; AAA24244.1; -.
EMBL; AE000195; AAC74015.1; -.
EMBL; D90730; BAA35675.1; -.
PMBL; D90731; BAA35681.1; -.
PIR; A03431; MMECF.
PIR; A25029; A25029.
                                                                                                                                                                                                                                                                                      SEQUENCE OF 23-34 AND 39-47.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nature 358:727-733(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 23-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               poring
    RALLAR RA
```

```
InterPro; IPR003229; OM_2.
InterPro; IPR001702; Porin_gram-ve.
InterPro; IPR001702; Porin_gram-ve.
InterPro; IPR00182; Econorins; 1.
PRINTS; PR00182; ECOLINIPORIN.
PROSITE; PS00576; GRAM_NEG_PORIN; 1.
Outer membrane; Transmembrane; Porin; Phage recognition; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 362;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                39333 MW; 3F0974D96DB65464 CRC64;
                                                                                                                                                                                                                                                                                                                                                      PROTEIN F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 1805; DB 1;
Pred. No. 2.9e-123;
                                                                                                                                                                                                                                                                                                                                                                      E (IN REF. 5).
G (IN REF. 5).
L (IN REF. 5).
                                                                                                                                                                                                                                                                                                                                                        OUTER MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                        ^ ^ ^
                                                                                                                                                                                                                                                                                                                       Complete proteome.
                                                                                                                                                                             ECO2DBASE; B036.0; 6TH EDITION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99.8%; 5
                                                                                                                                         PDB; 1BT9; 13-JAN-99.
SWISS-2DPAGE; P02931; COLI
                                                                                                                                                                                                                                                                                                                                    25 2 2 3 3 5 2 3 3 6 2 3 3 6 2 3 3 6 2 3 3 6 2 3 3 6 2 3 3 6 2 3 3 6 3 3 6 3 3 6 3 3 3 6 3 3 3 6 3 3 6 3 3 6 3 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6 3 6
                                                                                                                                                                                              EcoGene; EG10671; ompF
07-DEC-95.
07-FEB-95.
07-DEC-96.
                                                    07-DEC-96.
07-DEC-96.
07-DEC-96.
07-DEC-96.
07-DEC-96.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          291
306
329
329
351
353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                      3D-structure;
                                                                                                         1GFQ;
1MPF;
                                                                                        1GFP;
                                       1GFM,
                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TURN
TURN
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
HELIX
TURN
STRAND
STRAND
TURN
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                             TURN
                                                      DB;
```

m

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-K12 / W1110;
STRAIN-K12 / W1110;
Takemcoo K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
Mizuno T., Makino K., Nakata A., Yura T., Sampel G., Mizobuchi K.;
"Systematic sequencing of the Escherichia coli genome: analysis of the
"Systematic sequencing of the Escherichia coli genome: analysis of the
4.0 - 6.0 min (189,987 - 281,415bp) region.";
Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                  WEYNFOGNNSEGADAOTGNKTRLAFAGLKYADVGSFDYGRNYGVYYDALGYTDMLPEFGG 120
                                                                                                                                                                                                    180
                                                                                                                                                                                                                                         202
                                                                                                                                                                                                                                                                                                    WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 142
                                                                                                                                                                                                                                                                                                                                                              TNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT 300
  Gaps
                                       9
                                                                             82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
STRAIN-K12 / MG1655;
MEDLINE-97426617; PubMed-9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     й
Э
                                                            AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                  DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY
                                                                                                                                                                                                                                                                               EYEGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.
Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             for
of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Overbeeke N., Bergmans H., van Mansfeld F., Lugtenberg B.; "Complete nucleotide sequence of phoE, the structural gene phosphate limitation inducible outer membrane pore protein Escherichia coll X12."
J. Mol. Biol. 163:513-532(1983).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).
  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                          YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                              323 YYENKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 362
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
PHOE OR OMPE OR B0241.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        351 AA.
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-83189086; PubMed-6341601;
;
0
Matches 340; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PHOE_ECOLI
P02932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-K12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 2
PHOE_ECOLI
                                                                                                                                                          83
                                                                                                                                                                                                  121
                                                                                                                                                                                                                                       143
                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                       203
                                                                                                                                                                                                                                                                                                                                                            241
                                                                                                                                                                                                                                                                                                                                                                                                                                          301
                                                                                                                  ò
                                                                                                                                                        g
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCOORD BY A REPARED BY A REPARCE BY A REPAR
                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                MUTAGENESIS OF PHE-351.
MEDLINE=91162638; PubMed=1848301;
Struyve M., Moons M., Tommassen J.;
Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein.";
J. Mol. Biol. 218:141-148(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -i- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIANE-91172301; PubMed-1848682; Jap B.K., Wallan P.J., Gehring K.; Jap B.K., Wallan P.J., Gehring K.; Structural architecture of an outer membrane channel as determined by electron crystallography.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Struyve M., Visser J., Adriaanse H., Benz R., Tommassen J.;
"Topology of PhoE porin: the 'eyelet' region.";
Mol. Microbiol. 7:131-140(1993).
-1- FUNCTION: THIS IS ONE OF THE PROTEINS INDUCED WHEN E.COLI CELLS
ARE GROWN UNDER PHOSPHATE LIMITATION. ITS PROTEIN PORE IS
PARTICULARLY EFFICIENT IN THE UPTAKE OF INDREANIC PHOSPHATE,
PHOSPHORYLATED COMPOUNDS, AND SOME OTHER NEGATIVELY CHARGED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0011
                                                                                                                                                                              DNA and protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               R EMBL; AB000132; AA73345.1; -
R EMBL; D83536; BAA77910.1; -
R EMBL; D83536; BAA77910.1; -
R EMBL; N0786; CAA25362.1; -
R PIR; A03432; MMECPE.
R PDB; 1PHO; 31-OCT-93.
R SYLSS-2DPAGE; P02922; CCLI.
R ECOZDBASE; B037.0; 6TH EDITION.
R PROVED: PRO01729; PADE.
R INTERPRO; IPRO01729; PADE.
R INTERPRO; IPRO01729; POTID, STANDA, POSITE; PSO0076; GRAM-VEC_PORIN; 1.
R PRINTS; PRO0182; ECOLNEIDORIN; 1.
R PROSITE; PSO00776; GRAM-NEC_PORIN; 1.
R OUTER membrane; Transmembrane; Porin; Signal; 3D-structure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .
ш
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OUTER MEMBRANE PORE PROTEIN E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
MEDLINE-92375189; PubMed-1380671;
Cowan S. W., Schirmer T., Rummel G., Steiert M., Ghosh R.,
Pauptit R.A., Jansonius J.N., Rosenbusch J.P.;
Crystal structures explain functional properties of two I
Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                         Deutch A.H., Rushlow K.E., Smith C.J.; "Analysis of the Escherichia coli proBA locus by
                                                                                                                                                                                                                sequencing.";
Nucleic Acids Res. 12:6337-6355(1984).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X-RAY CRYSTALLOGRAPHY (6.0 ANGSTROMS).
                                     [5]
SEQUENCE OF 1-22 FROM N.A.
MEDLINE-84297232; Pubmed=6089111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-93172954; PubMed-7679770;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; V00316; CAA23605.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nature 358:727-733(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBUNIT: HOMOTRIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21
351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOLUTES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          porins.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
```

362 AA

```
Nature 413:852-856(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 59.6
Best Local Similarity 62.3
Matches 218; Conservative
      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        362
157
258
                                                                                        Salmonella typhimurium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   157
258
262 AA;
                                                                                                                            NCBI_TaxID=602;
                                                                                                                                                                                                                                                                                                               STRAIN-SL1303;
                             01-OCT-1994
     OMPD_SALTY
P37592;
                                                                                                                Salmonella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
     HDD BEAR AND DEAR AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F->Y, N, S, V: LESS RESISTANT TO TRYPSIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                180 YEYEG--FGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 238 TPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 351;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             38922 MW; 249E2E362C030C9A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61.4%; Score 1110; DB 1;
62.1%; Pred. No. 3.5e-73;
11ve 38; Mismatches 76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
  351
26
35
45
45
65
65
88
88
104
107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           351 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
  Best Local Sin
Matches 213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
MUTAGEN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
                                                                                               STRAND
TURN
STRAND
TURN
HELIX
                                                                                                                                                         TURN
TURN
TURN
STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                    STRAND
STRAND
TURN
STRAND
HELIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                    STRAND
STRAND
STRAND
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      URN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        309
 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseélsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                           MCClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nauyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K., Wilson R.K., "Complete genome sequence of Salmonella enterica serovar Typhimurium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 WEYNFOGN--NSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEF 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22 AEVYNKDGNKLDLYGKVHAQHYFSDDN----GSDGDKTYARLGFKGETQINDQLTGFGG 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-95011654; pubMed-7926834;
Hongo E., Morimyo M., Mita K., Machida I., Hama-Inaba H., Tsuji H.,
Ichimura S., Noda Y.;
"The methyl viologen-resistance-encoding gene smvA of Salmonella
typhimurlum.";
Gene 148:173-174(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                     Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Stydene; SG10249; ompD.
InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porth_gram.ve.
Pfam; PF00267; Gram.ve_portins; 2.
PROSITE; PS00576; GRAM_NEG_PORIN; 1.
Outer membrane; Transmembrane; Portin; Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OUTER MEMBRANE PORIN PROTEIN OMPD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N -> K (IN REF. 2).
MISSING (IN REF. 2).
9F68A9A7B948174C CRC64;
01-OCT-1994 (Rel. 30, Created)
15-JUN-2002 (Rel. 41, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
                                                                                         Outer membrane porin protein ompD precursor OMPD OR NMPC OR STM1572.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
STRAIN-LT2 / SGSC1412 / ATCC 700720;
MEDLINE-21534948; Pubmed-11677609;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AE008768; AAL20490.1; -.
EMBL; D26057; BAA05056.1; -.
HSSP; P02931; 1GFN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   39679 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 80-362 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
```

;

RESULT 3
OMPD_SALTY

S

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                       178 ISYEYEGFGIVGAYGAADRTNLQEAQPLGN----GKKAEQWATGLKYDANNIYLAANYGE 233
                                                                                                                                               234 TRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVN 293
GCDT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGS 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane. SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1; Length 350;
                                                                                                                                                                                                                                                                                                                                              294 YFEVGATYYFNKNMSTYVDYIINQIDSDN---KLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                            350 OUTER MEMBRANE PORE PROTEIN E. 38511 MW; 101665FEECC3AFA4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Interform, proposition of the pr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        047490;
15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
Outer membrane pore protein E precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         59.3%; Score 1072.5;
60.3%; Pred. No. 1.8e
:ive 40; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M28296; AAA24809.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBUNIT: HOMOTRIMER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Enterobacter cloacae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      350 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=550;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Enterobacter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PHOE_ENTCL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loca
Matches
119
                                                                                                                                                                                                                                                                                  g
                                                                                                              ð
                                                                                                                                                                 g
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                           ò
```

ŝ

Gaps

1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60

Conservative

207;

ô

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10;
                                                           Caps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 AEVYNKNGNKLDVYGKVKAMHYISDDDTK----DGDQTYVRFGFKGETQINDQLTGYGR
                                                                                                   121 D-TAYSDDFFVGRVGCVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS
                                           61. WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG
                                                                                                                                                              180 YEYEG--FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA
                                                                                                                                                                                                                       238 TPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59.3%; Score 1072; DB 1; Length 374; 60.1%; Pred. No. 2.1e-70; tive 38; Mismatches 81; Indels 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
OUTER MEMBRANE PROTEIN F.
BF821D385E441049 CRC64;
                                                                                                                                                                                                                                                                               298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                               308 GATYYFNKNMSAFVDYKINQIDDDNKLGVSSDDIVAVGMTYQF
                                                                                                                                                                                                                                                                                                                                                                                                               15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
Outer membrane protein F precursor (Porin ompF).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
Pfam; PR00267; Gram-ve_Porins; 1.
PRINTS; PR00182; ECCLNEIPORIN
PROSITE; PS00576; GRAM.NEG_PORIN; 1.
Outer membrane; Transmembrane; Porin; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; U81967; AAB69103.1; ALT_INIT.
HSSP; Q48473; 10SM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-UOC-51;
MEDLINE-97419518; PubMed-9274033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 41184 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Serratia marcescens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 374 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 215; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=615;
                                                                                                                                                                                                                                                                                                                                                                                  OMPF_SERMA
033980:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Serratia
                                                                                                                                                                                                                                                                                                                                                                     OMPF_SERMA
                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                      8
                                                                                                     à
                                                                                                                                 ద
                                                                                                                                                              ò
                                                                                                                                                                                         g
                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                           ద
```

351 AA; 38807 MW; 4E9C144070381B5A CRC64;

```
SEQUENCE
     So
                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@license.).
                                                                                                                                                                                                                                                                                               ISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRN 236
                                                                                                                                                                                                                                                                                                                           WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                           DT-AXSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERD--TARRSNGDGVGGS 177
                                                                                                                                                                                                                                                DIYIYSDNFWIGRINGVATYRNNNFFGLVDGLNFALQYQGKNQNDGRDVKKQNGDGWGIS 196
                                                                                                                                                                                                                                                                                                                                                                                                237 ATPI-TNKFTNT-----SGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAK--DVEGI 286
                                                                                                                                                                                                                                                                                                                                                                                                                            9
                            -1- SUBUNIT: HOMOTRIMER.
-1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
-1- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               G-DVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDN----KLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-93146376; PubMed-1337052;

MEDLINE-93146376; PubMed-1337052;

MEDLINE-93146376; PubMed-1337052;

MEDLINE-93146376; PubMed-1337052;

Characterization of the Citrobacter freundii phoE gene and development of C. freundii-specific oligonuclectides.";

FEMS Microbiol. Lett. 78:139-204 (1992).

-I- FUNCTION: THIS IS ONE OF THE PROTEINS INDUCED WHEN CELLS ARE GROWN UNDER PHOSPHATE LIMITATION. ITS PROTEIN PORE IS PARTICULARLY EFFICIENT IN THE UPTAKE OF INORGANIC PHOSPHATE, PHOSPHORYLATED COMPOUNDS, AND SOME OTHER NEGATIVELY CHARGED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Citrobacter.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OUTER MEMBRANE PORE PROTEIN E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HISSP, P02932; 1PHO.
InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
Pfam; PF00267; Gram-ve_porins; 1.
PRINTS; PR00182; ECOLNEIPORIN.
PROSITE; PS00576; GRAM_MEG_PORIN; 1.
Outer membrane; Transmembrane; Porin; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        351 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Outer membrane pore protein E precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; X68021; CAA48162.1; -. PIR; S25520; S25520.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Citrobacter freundil
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=546;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PHOE_CITFR
Q01605;
                                                                                                                                                                                                                                              137
                                                                                                                                                                                                                                                                                                                                               197
                                                                                                 61
                                                                                                                                             78
                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               287
                                                g
                                                                                                                                             qq
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HE WELL HER WALLE OF THE PROPERTY OF THE PROPE
ò
                                                                                                 ð
                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                               à
```

```
Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M., Baker S., Basham D., Brooks K., Chillingworth T., Connerton P., Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J., Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K., Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C., Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                              180 YEYEG--FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         238 TPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV 297
                                                              Gaps
                                                                                                                   9
                                                                                                                                                22 AEVYNKNGNKLDLYGKVKAMHYMTDYDSK----DGDQSYIRLGFKGETQINDELTGYGR 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane. SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18."; Nature 413:848-852(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -!- FUNCTION: THIS IS ONE OF THE PROTEINS INDUCED WHEN CELLS ARE GROWN UNDER PHOSPHATE LIMITATION. ITS PROTEIN PORE IS PARTICULARLY EFFICIENT IN THE UPTAKE OF INORGANIC PHOSPHATE, PHOSPHORYLATED COMPOUNDS, AND SOME OTHER NEGATIVELY CHARGED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                           16:
     351;
                                                           Indels
     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=IMSS-1;
Torres A., Puente J.L., Calva E.;
Submitted (AUG-1993) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       309 GATYYENKNMSAFVDYKINQLDSDNKLGINNDDIVAVGMVYQF 351
  59.1%; Score 1070; DB 1; 59.2%; Pred. No. 2.7e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-JUN-2002 (Rel. 41, Last sequence update) 15-JUN-2002 (Rel. 41, Last annotation update) Outer membrane pore protein E precursor.
                                                        48; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-CT18;
MEDLINE-21534947; Pubmed-11677608;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (Rel. 35, Created)
                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOLUTES.
SUBUNIT: HOMOTRIMER.
                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Salmonella typhi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Salmonella.
NCBI_TaxID=601;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PHOE OR STY0365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1997
                                                        Matches 203;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PHOE_SALTI
Q56119;
Query Match
                                   Local
                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                      11
```

~

```
유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseelsb-sib.ch).
                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                               297
                                                                                                                                                                                                                                                                                                                                      WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                               D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                                                                                      YEYEG---FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    307
                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                       1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                        238 TPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Salmonella typhimurium.
Bacteria: Proteobacteria: gamma subdivision; Enterobacteriaceae;
Salmonella.
                                                                                                  HSSP; P02932; 1PHO.
InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
Pfam; PF00267; Gram-ve.
Promis; PR00182; ECCINEIPORIN; 1.
PROSITE; PS00576; GRAM_NEG_PORIN; 1.
Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-93083994; PubMed-1280609;
Splerings G., Elders R., van Lith B., Hofstra H., Tommassen J.;
"Characterization of the Salmonella typhimurium phoE gene and
development of Salmonella specific DNA probes.";
                                                                                                                                                                                                                                                                          16;
                                                                                                                                                                                                                                                   58.5%; Score 1058; DB 1; Length 350; 58.0%; Pred. No. 2e-69;
                                                                                                                                                                           BY SIMILARITY.
OUTER MEMBRANE PORE PROTEIN E.
                                                                                                                                                                                                                                                                         82; Indels
                                                                                                                                                                                                                              890F5F8C5EF0CC77 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E -> K (IN REF. 1).
F -> G (IN REF. 1).
A -> G (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF
                                                                                                                                                                                                                                                              ; Pred. No. 2e-69; 46; Mismatches 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
PHOE OR STM0320.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                           20 BY
350 OU
63 E
80 F
229 A
38744 MW;
                                                                               EMBL; X74595; CAA52672.1; -.
EMBL; AL627266; CAD08790.1;
                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                     21 3
63
80
229
350 AA;
                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=602;
                                                                                                                                                                                                                                                                       Matches 199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PHOE_SALTY P30705;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-LTZ;
                                                                                                                                                                                                        CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                Local
                                                                                                                                                                             SIGNAL
                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PHOE_SALTY
                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                      180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                 윱
                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL outstation
                                                                                                                between the Swiss Institute of Bioinformatics and the EMBL outstati. the Buropean Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YEYEG--FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV 297
                                                                                                                                                                                                                                                                                                                                                                                                                               -:- SUBUNIT: HOMOTRIMER.
-:- SUBCELLULAR LOCATION: Integral membrane protein, Outer membrane.
-:- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       82; Indels 16; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nature 413:852-856(2001).
-1- FUNCTION: THIS IS ONE OF THE PROTEINS INDUCED WHEN CELLS ARE GROWN UNDER PHOSPHATE LIMITATION. ITS PROTEIN PORE IS PARTICULARLY EFFICIENT IN THE UPTAKE OF INORGANIC PHOSPHATE. PHOSPHORYLATED COMPOUNDS, AND SOME OTHER NEGATIVELY CHARGED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     58.5%; Score 1058; DB 1; Length 350; 58.0%; Pred. No. 2e-69; 1.1ve 46; Mismatches 82; Indels 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Εij.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OUTER MEMBRANE PORE PROTEIN 9D1EE3355AF59877 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; X68023; CAA48164.1; -.
EMBL; AE008709; AAL19276.1; -.
PIR; S25525; S25525.
HSSP, P02932; 1PH0.
StyGene; SG10291; PhoE.
InterPro; IPR001229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
PRINTS; PR00182; ECOLNEIPORIN;
PROSITE; PS00576; GRAM_NEG_PORIN; I.
                                                                   STRAIN-LT2 / SGSC1412 / ATCC 700720;
MEDLINE-21534948; PubMed-11677609;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE 350 AA; 38762 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 58.0%
Matches 199; Conservative
Gene 122:45-52(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    308
```

us-09-490-291-9.rsp

```
õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                  MEDLINE-83213433; PubMed-6304064; Mizuno I., Chou M.Y., Inouye M.; "A comparative study on the genes for three porins of the Escherichia coll outer membrane. DNA sequence of the osmoregulated ompC gene."; J. Biol. Chem. 258:6932-6940(1983).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nogami T., Mizuno T., Mizushima S.; "Construction of a series of ompF-ompC chimeric genes by in vivo homologous recombination in Escherichia coli and characterization of the translational products."; "Bacteriol. 164:797-801(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN=K12 / EMG2;
MEDLINE-97443975; PubMed-9298646;
Link A.J., Robison K., Church G.M.;
"Comparing the predicted and observed properties of proteins encoded
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 1-22 FROM N.A., AND SEQUENCE OF 22-40.
MEDILTRE-8431325; PubMed-6297988;
MAIZUNO T., Chou M.-Y., Innouye M.;
"DNA sequence of the promoter region of the ompC gene and the amino acid sequence of the signal peptide of pro-OmpC protein of Escherichia col.,";
FEBS Lett. 151:159-164(1983).
                                                                                                                                                                                                                                                                               Care From N.A.
STRAIN-KIZ / MG1655,
MEDLINE-97426617; PubMed-9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V. Kiley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN~K12 / BHB2600;
Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
                                                                                                                                      Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Itch T., Alba H., Babba T., Fujita K., Hayashi K., Inada T., Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T., Mizbuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C., Yamamoto Y., Horluchi T.; "A 460-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 40.1-50.0 min region on the linkage map.";
                                                                        01-APR-1988 (Rel. 07, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Outer membrane protein C precursor (Porin ompC) (Outer membrane
                                                                                                                                                                                                                                                                                                                                                                                     "The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
                                    367 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-97251358; PubMed-9097040;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 32-57 FROM N.A. MEDLINE-86033642; PubMed-2997131;
                                                            (Rel. 07, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 218-367 FROM N.A.
                                                                                                                         OMPC OR MEOA OR PAR OR B2215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA Res. 3:379-392(1996).
                                   STANDARD;
                                                                       (Rel. 07,
                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SECUENCE OF 22-30
                                                                                                                                                                            NCBI_TaxID=562;
                                                                                                              protein 1B).
                                                           01-APR-1988
01-APR-1988
                                 OMPC_ECOLI
P06996;
                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-K12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Church G.
        RESULT 9
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ilensib.ch).
                                                                                                                                                                                                                                                 separation using two-dimensional gel electrophoresis..;
Electrophoresis.19:87-844(1998).
-!- FUNCTION: FORM PASSIVE DIFFUSION PORES WHICH ALLOW SMALL MOLECULAR
WEIGHT HYDROPHILLIC MATERIALS ACROSS THE OUTER MEMBRANE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -1- SUBUNIT: HOMOTRIMER.
-1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
-1- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9/
                                                                                                                               MEDLINE-98291876; PubMed-9629924;
Molloy M.P., Herbert B.R., Walsh B.J., Tyler M.I., Traini M.,
Sanchez J.-C., Hochstrasser D.F., Williams K.L., Gooley A.A.;
"Extraction of membrane proteins by differential solubilization for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22 AEVYNKDGNKLDLYGKVDGLHYFS-----DNKDVDGDQTYMRLGFKGETQVTDQLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     226 YLAANYGETRNATPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            167 RRSNGDGVGGSISYEYEGFGIVGAYGAADRTNLQE-AQPLGNGKKAEQWATGLKYDANNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00576; GRAM_NEG_PORIN; 1.
Outer membrane; Transmembrane; Porin; Phage recognition; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         58.1%; Score 1051; DB 1; Length 367; 58.9%; Pred. No. 6.6e-69; Indels 34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22 367 OUTER MEMBRANE PROTEIN C. 367 AA; 40368 MW; 6A49370CC8A1A225 CRC64;
the genome of Escherichia coli K-12."; lectrophoresis 18:1259-1313(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
Pfam; PF00267; Gram-ve_porins; 1.
PRINTS; PR00182; ECOLNEIPORIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; A20867; MMECPC.
PIR; B25029; B25029.
PIR; A18885; A18885.
HSSP; O52503; 111V.
ECO2DBASE; A035.5; 6TH EDITION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AE000310; AAC75275.1; -. EMBL; D90850; BAA15998.1; -. EMBL; U00008; AAA16412.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; K00541; AAA24243.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 58.9
Matches 212; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21
367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OmpC
                                                                                                              STRAIN-K12 / W3110;
                                                                              SEQUENCE OF 22-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EcoGene; EG10670;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

.; 6

```
MEDLINE-86304457; PubMed-3017988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     40290 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; J02580; AAA32301.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; dsDNA viruses, no
unclassified Siphoviridae.
NCBI_TaxID=10738;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .. <u>-:</u> ___
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                365 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                       PA-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 210;
                                                                                                                                                                                                                        PORI_BPPA2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                     PORI_BPPA2
                                                                  à
                                                                                                        원
                                                                                                                                                                                                                                                  NOTE THE REPRESENTATION OF THE PROPERTY OF THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseelib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180 .YEYEG--FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     195 YDFGGTDFAVSAAYTSSDRTNDQNLLARAQGSKAEAWATGLKYDANNIYLATMYSETRKM 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       238 TPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22 AEVYNKNANKLDVYGKIKAMHYFSDYDSK-----GDQTYVRFGIKGETQINEDLIGYRR 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                   van der Ley P., Bekkers A., van Meersbergen J., Tommassen J.;

A comparative study on the phoE genes of three enterobacterial
species. Implications for structure-function relationships in a
pore-forming protein of the outer membrane.";

Eur. J. Biochem. 164:469-475(1987).

-! FUNCTION: THIS IS ONE OF THE PROTEINS INDUCED WHEN CELLS ARE
GROWN UNDER PHOSPHATE LIMITATION. ITS PROTEIN PORE IS
PHOSPHORYLATED COMPOUNDS, AND SOME OTHER NEGATIVELY CHARGED
                                                                                                                                                                                                                                                                  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Klebsiella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 351;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OUTER MEMBRANE PORE PROTEIN E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              86A5286C12502EC7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           58.0%; Score 1050; DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7.4e-69
                                                                                                                                                01-APR-1993 (Rel. 25, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Signal.
                                                                                     351 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pram; PF00267; Gram-ve_porins; 1..
PRINTS; PR00182; ECOLNEIPORIN; 1.
PROSITE; PS00576; GRAM_NEG_PORIN; 1.
Outer membrane; Transmembrane; Porin; Signs SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     43; Mismatches
                                                                                                                                                                                                  Outer membrane pore protein E precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
                                                                                     PRT:
                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-87190417; PubMed-3032618;
                                                                                                                                (Rel. 25, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38894 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M28295; AAA25121.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                58.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 201; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBUNIT: HOMOTRIMER.
                                                                                                                                                                                                                                             Klebsiella pneumoniae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           351 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                NCBI_TaxID=573;
                                                                                                                           01-APR-1993
                                                                                   PHOE_KLEPN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local
                                                           PHOE_KLEPN
                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lsb-sib.ch/announce.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Jambdold bacterlophage.";
J. Blol. Chem. 261:12723-12732(1986).
J. Blol. Chem. 261:1273-12732(1986).
J. GRANGELIONS ARE MAJOR PROTEINS FOUND IN THE OUTER WEMBRANES
THAN 4000-6000 DALTONS).
J. SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (By SIMILARITY).
-!- MISCELLANEOUS: A PORIN GENE CAN ALSO BE FOUND IN THE GENOMES OF CERTAIN LAMBDOID BACTERIOPHAGE, AND ITS PROTEIN IS EXPRESSED IN THE LYSOGENIC STATE. IN E.COLI THE EXPRESSION OF OMPC AND OMPF PROTEINS IS THEN REDUCED SUBSTANTIALLY.
-!- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        no RNA stage; Caudovirales; Siphoviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Blasband A.J., Marcotte W.R. Jr., Schnaitman C.A.; "Structure of the lc and nmpC outer membrane porin protein genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OUTER MEMBRANE PORIN PROTEIN LC.
H -> R (IN REF. 1; AA SEQUENCE).
OFBC0531FB9C0205 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            84; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               340
                                    309 GLTYYFNKNMNAFVDYKINQLKSDNKLGINDDDIVALCMTYQF
298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              57.8%; Score 1046; DB 1; 59.5%; Pred. No. 1.5e-68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIRI, D25647; MMBPP2.
HSSP, 052503; 111V.
InterPro; IPR0013229; OMP_2.
Pfam; PF00267; Gram-ve_porins; 1.
PRINTS; PR00182; ECOLNEIPORIN.
PRINTS; PR00182; ECOLNEIPORIN.
OUTGIT membrane; Transmembrane; Porin; Signal.
                                                                                                                                                                                                                                                                                                                  01-APR-1988 (Rel. 07, Created)
01-APR-1988 (Rel. 07, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                          Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                 Outer membrane porin protein LC precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            or send an email to license@isb-sib.ch)
```

```
STITE BORRER BOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          <u>،</u>
WEYEFKGNRAE-SQGSSKDKTHLAFAGLKFGDYGSIDYGRNYGVAYDIGAWTDVLPEFGG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ρλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Coll J.L., Heyde M., Portaller R.,
"Expression of the nmpC gene of Escherichia coli K-12 is modulated external pH. Identification of cis-acting regulatory sequences involved in this regulation.";
                                                                                                          DT-AYSDDFFVGRVGGVATYRNSNFFCLVDGLNFAVQYLGKNER---DTARRSNGDGVGG
                                                                                                                                                                                                      SISYEYEGFGIVGAYGAADRINLQE-----AQPLGNGKKAEQWATGLKYDANNIYLAAN
                                                                                                                                                                                                                                                                                                                                             YGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=96196428; PubMed=8648624;
Mahdi A.A., Sharples G.J., Mandal T.N., Lloyd R.G.;
Halliday junction resolvases encoded by homologous rusA genes in
Escherichia coli K.12 and phage 82.";
J. Mol. Biol. 257:561-573(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-86304457; PubMed-3017988; Blasband A.J., Marcotte W.R. Jr., Schnaitman C.A.; Sztructure of the lc and nmpC outer membrane porin protein genes lambdoid bacteriophage...; J. Biol. Chem. 261:12723-12732(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
STAIN-KIZ / MG1655;
MEDLINE-97426617; Pubmed-9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland v Blattner F.R., Glundo-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., L.
Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LVNYFEVGATYYFNKNMSTYVDYIINQIDSDN----KLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "The complete genome sequence of Escherichia coll K-12."; Science 277:1453-1474(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME ECOLI STANDARD; PRT; 365 AA. P21420; P77189; 01-MAY-1991 (Rel. 18, Created) 01-MAY-1991 (Rel. 18, Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update) outer membrane porin protein nmpC precursor. NMPC OR PHMA OR B0553.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-JL5502;
MEDLINE-94335635; Pubmed-8057841;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .. Microbiol. 12:83-93(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 347-365 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 1-17 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
STRAIN-Mutant CS384;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                             258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    291
79
                                                                                                                                   138
                                                                                                                                                                                                      177
                                                                                                                                                                                                                                                                          198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
硆
                                                                   ð
                                                                                                                                   g
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HDD DTTC DDTTC DDTTC DDTTC DDTTC DDTTC DDTTC DDTTC DDTTC DDTTC DTTC DTTC DDTTC DTTC DT
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib.ch).
SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane. MISCELLANBOUS: IN WILD-TYPE STRAINS OF E.COLI K12, THE NMPC OPEN READING FRAME IS INTERRUPTED BY AN ISS INSERTION AND GENERATES A HYBRID OPEN READING FRAME TAND TO STRAIN OPEN READING FRAME THAT IS NOT EXPRESSED. HOWEVER, IN MUTANT STRAIN CS348, THE ISS ELEMENT HAS BEEN DELETED AND NMPC IS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNER---DTARRSNGDGVGG 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24 AEIXNKDSNKLDLYGKVNAKHYFSSNDAD-----DGDTIYARLGFRGETQINDQLTGFGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SISYEYEGFGIVGAYGAADRTNLQE-----AQPLGNGKKAEQWATGLKYDANNIYLAAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      YGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                258 YSETQNMTVFADHFV-----ANKAQNFEAVAQYQFDFGLRPSVAYLQSKGKDLGVWGQQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00267; Gram-ve_porins; 1.
PRINTS; PR00182; ECOLNEIPORIN.
PROSITE; PS00576; GRAM_NEG_PORIN; 1.
Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OUTER MEMBRANE PORIN PROTEIN NMPC.
K -> N (IN REF. 2 AND 3).
6E5128D4847FB4F8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             57.6%; Score 1042; DB 1; Length 365; 59.5%; Pred. No. 2.9e-68; ive 35; Mismatches 84; Indels 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LVNYFEVGATYYFNKNMSTYVDYIINQIDSDN---KLGVGSDDTVAVGIVYQF
                                                                                                                                                                        SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
PHOE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M13457; AAA23728.1; ALT_SEQ.
EMBL; AE000160; AAC73654.1; ALT_SEQ.
EMBL; U82598; AAR40749.1; ALT_SEQ.
EMBL; 23542; CAA84594.1; -:
EMBL; X92587; CAA63325.1; -:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; A25647; MMCCMC.
HSSP; Q48473; 10SM.
ECOGENE; EG10659; mmC.
InterPro; IPR0013229; OMP_2.
InterPro; IPR001702; PorIn_gram-ve.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24 365 C
326 326 R
365 AA; 40316 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 59.5
Matches 210; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 13
PHOE_KLEOX
ID PHOE_KLEOX
AC Q01606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                291
```

.;

```
P77747; P76854;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIGNAL
  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121 D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   180 YEYEG--FGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane. SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              238 TPITHKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                    Spierings G., Ockhuijsen C., Hofstra H., Tommassen J.;
"Characterization of the Citrobacter freundii phoE gene and
development of C. freundii-specific oligonucleotides.";
FEMS Microbiol. Lett. 78:199-204 (1992).
-!- FUNCTION: THIS IS ONE OF THE PROTEINS INDUCED WHEN CELLS ARE
GROWN UNDER PHOSPHATE LIMITATION. ITS PROTEIN PORE IS
PARTICULARLY EFFICIENT IN THE UPTARE OF INORGANIC PHOSPHATE,
PHOSPHORYLATED COMPOUNDS, AND SOME OTHER NEGATIVELY CHARGED
Klebsiella oxytoca.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18;
                                                                                                                                                                                                                                                                                                                                                                                                                                 OUTER MEMBRANE PORE PROTEIN E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                            D0B0A5427542B3F5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      57.5%; Score 1041; DB 1; 58.6%; Pred. No. 3.3e-68;
                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF00267; Gram-ve_porins; 1.
PRINTS; PR00182; ECOLNEIPORIN.
PROSITE; PS00576; GRAM_NEG_PORIN; 1.
Outer membrane; Transmembrane; Porin; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        43; Mismatches
                                                                                                                                                                                                                                                                                                                                                InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
                                                                            MEDLINE-93146376; PubMed-1337052;
                                                                                                                                                                                                                                                                                                                                                                                                                                           38550 MW;
                                                                                                                                                                                                                                                                                                                             EMBL; X68022; CAA48163.1; -. HSSP; P02932; 1PHO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 201; Conservative
                                                                                                                                                                                        SUBUNIT: HOMOTRIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                22 3
349 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                       SEQUENCE FROM N.A.
                                NCBI_TaxID=571;
                                                                                                                                                                                SOLUTES
                       Klebsiella
                                                                 STRAIN-K26
                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                       SIGNAL
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
```

¥.

377

PRT;

OMPN_ECOLI

RESULT 14 OMPN_ECOLI ID OMPN_I

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lsb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M., Makin T., Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N., Sampel G., Seki Y., Sivasundaram S., Tagami H., Takeda J., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T., A. A. Tagami H., Takeda J., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T., Corresponding to the 28 0-40.1 min region on the linkage map.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        78; Indels 38; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [3] SEQUENCE FROM N.A., SEQUENCE OF 22-33, AND CHARACTERIZATION. MEDIINE-98317278; Pubmed-9642192; Prilipov A., Phale P.S., Koebnik R., Widmer C., Rosenbusch J.P.; "Identification and characterization of two quiescent porin genes,"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SECUENCE FROM N.A.
STRAIN-M.L. / MG1655;
MEDI.INE-97426617; Pubmed-9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V. Alley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AE000234; AAC74459.1; -.
EMBL; D90775; BAA14981.1; -.
EMBL; D90775; BAA14986.1; -.
HSSP; P02931; GD931; GD90775; D0007.

ECGGene; EG13375; OMP_2.
InterPro; IPR003229; OMP_2.
PICAPTO; IPR01702; POTING_gram-ve.
Pfan; PF00267; GTam-ve_poring; 1.
PRINTS; PR00182; ECOLNEIPORIN.
Outer_membrane; Transmembrane; Porin; Signal; Complete proteome.
                                                                                                                                                               Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1; Length 377;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4420D2CB9BC10F95 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OUTER MEMBRANE PROTEIN N.
15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
16-GCT-2001 (Rel. 40, Last annotation update)
Outer membrane protein N precursor (Porin ompN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 57.2%; Score 1035; DB 1. Best Local Similarity 56.9%; Pred. No. 9.7e-68 Matches 209; Conservative 42; Mismatches 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-97251357; PubMed-9097039;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41220 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 377 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                  NCBI_TaxID-562;
                                                                                                                                    OMPN OR B1377
                                                                                                                                                                                                                            Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-K12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M., Baker S., Basham D., Brooks K., Chillingworth T., Connerton P., Croin A., Davis P., Davies R.M., Dowd L., White N., Farrar J., Aroth A., Larsen T.S., Learber S., Moule S., O'Gaora P., Parry C., A. Ouail M., Rutherford K., Simmonds M., Skelton J., Stevens K., Mitchead S., Barrell B.G.; Mule S., O'Gaora P., Parry C., Mhitchead S., Barrell B.G.; Mule S., O'Gaora P., Parry C., Mhitchead S., Barrell B.G.; Multiple drug resistant Salmonella archerica serovar Typhi CT18."; Mule S., O'Gaora P., Stevens K., Mule S., Complete genome sequence of a multiple drug resistant Salmonella archerica serovar Typhi CT18."; Mature 413:848-852(2001).

C. I. FONGTION: OMPF IS A PORIN THAT PASSIVE DIFFUSION PORES WHICH ALLOW SMALL MOLECULAR WEIGHT HYDROPHILIC MATERIALS ACROSS THE OUTER MEMBRANE. IT IS ALSO RECEPTOR FOR THE BACTERIOPHAGE T2 (BY SIMILARITY).

C. I. SUBURITY: HOMOTRIMER (BY SIMILARITY).

C. I. SUBURITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                              253
                                               DT-AYSDDFFVGRVGCVATYRNSNFFGLVDGLNFAVQYLGKNE------RDTAR 167
                                                                                                                                                             RSNGDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIY 226
                                                                                                                                                                                                                          LAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVE-- 284
                                                                                                                                                                                                                                                                                       -----GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVA 333
                                                                                                                                                                                                                                                                                                        22 AEVYNKDGNKLDLYGKVDGLHYFSDNSAK-----DGDQSYARLGFKGETQINDQLTGYGQ
                               WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OMPFOR OMESS OR STY1002.
Salmonella typhi.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-JUN-2002 (Rel. 41, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Outer membrane protein F precursor (Porin ompF) (Outer membrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-IMSS-1;
Fernandez-Mora M., Calva E.;
Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  363 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-21534947; PubMed-11677608;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Rel. 37, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                    334 VGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                     :|:||||
LGLVYQF 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Salmonella.
NCBI_TaxID=601;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-DEC-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-JUN-2002
15-JUN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-CT18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                OMPF_SALTI
Q56113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein S3)
                                                                                                                                                             168
                               61
                                                                                              121
                                                                                                                                                                                                                          227
                                                                                                                                                                                                                                                                                       285
                                                                                                                                                                                                                                                        254
                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11D
DD41DD41DD
DD41DD
g
                               ò
                                                              a
                                                                                              à
                                                                                                                          g
                                                                                                                                                           ò
                                                                                                                                                                                         엄
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                    à
```

```
and for commercial
               entities requires a license agreement (See http://www.isb-sib.ch/announce.or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                             .;
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23 AEIYNKDGNKLDLYGKAVGRHVWTT-TGDSK---NADQTYAQIGFKGETQINTDLTGFGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DT----AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRS-NGDGVGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         177 SISYEYEGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 237 ATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGI-GDVDLVNYF
                                                                                                                                       InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
InterPro; IPR001702; Porin_gram-ve.
Pfam; PR00182; ECOLNEIPORIN: 1.
PROSITE; PS00576; GRAM_REG_PORIN; 1.
Outer membrane; Transmembrane; Porin; Phage recognition; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                             13;
                                                                                                                                                                                                                                                                                                                                                                                                      DB 1; Length 363;
                                                                                                                                                                                                                                                                                                                                                                                                                                           80; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EVGATYYFNKNMSTYVDYIINQIDSDNKLG--VGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Usage by
                                                                                                                                                                                                                                                                                             22 BY SIMILARITY.
363 OUTER MEMBRANE PROTEIN F.
307 D - 6 (IN REF. 1).
40106 MW; F5059B37EA116859 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                             Query Match 57.0%; Score 1031.5; DB Best Local Similarity 57.6%; Pred. No. 1.7e-67 Matches 200; Conservative 54; Mismatches 8
modified and this statement is not removed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Search completed: December 18, 2002, 16:18:23 Job time : 6.23408 secs
                                                                              EMBL; X89757; CAA61905.1; -. EMBL; AL627268; CAD05399.1; HSSP; P02931; 1GFN.
                                                                                                                                                                                                                                                                                             22
363
307
                                                                                                                                                                                                                                                                                                                                     307 3
363 AA;
                                                                                                                                                                                                                                                                            Complete proteome.
                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
```

December 18, 2002, 16:13:04; Search time 20.0203 Seconds (without alignments) 3509.535 Million cell updates/sec GenCore version 5.1.3 Copyright (c) 1993 - 2002 Compugen Ltd. OM protein - protein search, using sw model Run on:

US-09-490-291-9 1809 1 AEIYNKDGNKVDLYGKAVGL......NKLGVGSDDTVAVGIVYQFA 341 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

671580 seqs, 206047115 residues Searched:

671580 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

SPTREMBL_21:*

1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_human:*
5: sp_nvertebrate:*
5: sp_nvertebrate:*
6: sp_phage:*
7: sp_phage:*
7: sp_phage:*
7: sp_vorganelle:*
7: sp_vorgane

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

		æ				
Result No.	Score	Query Match	Query Match Length DB	DB	ΙD	Description
-	1789		362	16	O8XDF1	O8xdfl escherichia
7	1390.5	•	351	~	Q93K98	093k98 enterobacte
٣	1113	61.5	351	16	Q8X7N5	08x7n5 escherichia
4	1082		371	16	082686	O8zgs6 versinia pe
Ŋ	1061	-	363	~	Q93K99	093k99 enterobacte
9	1044.5		374	~	087754	087754 klebstella
7	1036	57.3	377	7	085030	085030 escherichia
80	1035.5	57.2	372	7	Q93K39	Q93k39 klebsiella
6	1033	57.1	367	16	08XE41	O8xe41 escherichia
10	1029.5	56.9	364	7	Q9K597	09k597 escherichia
11	1029.5	56.9	378	7	Q9K3E6	09k3e6 salmonella
12	1028.5	56.9	366	7	Q9RH85	Q9rh85 escherichia
13	1022.5	56.5	333	N	Q9RAW3	Ograw3 klebsiella
14	1022.5	56.5	333	7	Q9RAW2	Ograw2 klebsiella
15	1022.5	56.5	363	~	Q9K3E7	09k3e7 salmonella
16	1021.5	56.5	333	~	O9RR59	Ogrr59 klebsiella

Q9s613 klebsiella O8vue7 klebsiella Q8vpa5 klebsiella Q9f889 salmonella O8zp14 salmonella	<b>~~~~</b>		clostridio salmonella klebsiella salmonella salmonella	09exh8 klebsiolla 092c71 yerstila pe 08x90 escherichia 08xb87 escherichia 052641 rahnella aq 08x96 escherichia 08x942 escherichia 08x942 escherichia
2 Q9S613 2 Q8VUE7 2 Q8VPA5 2 Q9F889 16 Q8ZPL4	9999	00	00000	2 Q9EXH8 2 Q9Z71 16 Q8X9Q0 16 Q8XBB7 2 Q5Z641 16 Q8XBB6 16 Q8X9Q2 2 P94857
333 333 360 377	375 375 375 375	375 376 360 398 374	3452 342 342 342 342	315 315 224 224 130 123
55 55 55 55 55 55 55 55 55 55 55 55 55	00000000000000000000000000000000000000	555.2 54.8 54.4 53.1	51.3 51.7 51.6 51.6 50.8	45.1 42.6 35.9 23.7 23.4 20.6
1019.5 1019.5 1019.5 1017.5	1007 1005 1000 999	998 995.5 992 984.5 960	9346 934.5 933 933	815 770 649.5 539.5 428.5 423 372.5
17 18 19 20 21	22 24 25 25 26	27 28 29 30 31	3 3 3 3 3 3 4 3 5 4 5 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	33 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3

#### ALIGNMENTS

RESULT 1

				Last sequence update)	Last annotation update)				Enterobacter1aceae;							Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,	Posfal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,	E.T., Potamousis K.,	3., Schwartz D.C.,		"Genome sequence of enterohaemorrhagic Escherichia coli 0157:H7.";						Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,	., Tanaka M., Tobe T.,	lida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,		"Complete genome sequence of enterohemorrhagic Escherichia coli	0157:H7 and genomic comparison with a laboratory strain K-12.";			
	Ş			in	lo O				io O							Жa	Ä	;	inte	ä		ž Č						BM8	83	:	8	rhac	Sorg			
	362 AA.			Suce	tati				V181							:	rego	ļu,	na 18	×.		ස් ප						Ska	urat	ບ	agav	ű	la			
			<del>G</del>	edae	nno				subdivision;					927		'n	<u>ნ</u>	out	급	<u>ب</u>		ag 1.						Κū	Ē,	kaw	htn	oher	ь			
	PRT;		Created)	Ť.	r B	(Ia, b, F).			S					700	51;	rla	ŝ	Φ,	Α.	Lin		rrh					96;	Ξ.	×.	asa	s S	ter	χt			
	Δ,		Cre	Las	Las	a,b			gamma					ည	3065	BG	38 P	s.	ET.			Ē					587	Ĭ,	/ama	·	ž	eu	gon		٠,	
			20,	20,	21,									STRAIN-0157:H7 / EDL933 / ATCC 700927;	MEDLINE-21074935; PubMed-11206551;	III'	Evar	11n)	:	E-1	•	roh				STRAIN=0157:H7 / RIMD 0509952;	Ë	hnis	aka)	E E	Kuhara S., Shiba T., Hattori M., Shinagawa H.;	e o	ari	-	1	7
	RY;					Outer membrane protein la	012	Escherichia coli 0157:H7.	Bacteria; Proteobacteria;					33	Med		.:	χ,	3 2	man	Welch R.A., Blattner F.R.;	inte	5			0.5	Med	0	z	puo	Hat	enc	duo.		EMBL: AP002553; BAB34435.1;	InterPro; IPR003229; OMP_2
	4INA		3Lre	3Lre	3Lre	tei	SCS1	1157	acte					501.9	Pub	i,	G.F	ח	139	ara	er	ž e	3 2 3			ZI ME	Pub	χ.	ei Ei	Ξ.	:	sedn	20	100	<b>VB34</b>	39;
	PRELIMINARY;		(TrEMBLrel.	(TrEMBLrel.	(TrEMBLrel.	pro	S. E	11	agos		34;		N.A.	_	35;	unke	hew	kett	Da	anth	at t	9	-533		Ä.	_	31;	kind	anpo	Ŧ	ba J	je s	E C	22	ă	0322
	PRI					ane	92	8	rot		8333		8	: H2	749	F	May	Hac	; :	An	B	neu	529		- Σ	: H2	562	Μæ	g ti	kam.	Shi	eno	ge	100	553	PR0
			01-MAR-2002	01-MAR-2002	01-JUN-2002	apr.	212	hia	ď.	Escherichia	NCBI_TaxID-83334;		SEQUENCE FROM N.A.	157	210	Τ.,		:	ы.	٦ :	Ä	sed			SEQUENCE FROM N.A.	157	211	Ξ,	:	Ta	:	ė.	and		000	
	덛	ij	R-2	\R-2	JN-2	E.	o R	ric	ria	ric	Tax		SNCE	Ž	NE-	ż	5.	11 6	šec	ca	ξ.	ЭЩС	ė.		SNCE	Ž	ENE-	'nį	<u>ب</u>	Ţ.	e S	let	Н7	es.	Z	Pro
ı	Q8XDF1	Q8XDF1;	]-X	1-M	J-1	uter	MPF	sche	acte	sche	CBI.	7	EOG	IRA	EDI	erna	<b>98</b> C	gge	roth	podg	e10	Serio	ater	2]	EQUE	FR	EDI	ayas	an	ida	uhaı	Comp	157	¥ 9	MBL	ntei
DF1	õ	õ	0	0	0	õ	ਠ	Щ	ã	ŭ	ž	<u>ٺ</u>	ß	ັດ	Ξ	ă.	×	ă,	G	¥	ž	_	ž	<u>ت</u>	ŝ	ò	Σ	Ĩ	Ĩ	H	ž	Ĕ	o i	3 6	ធី	A
O8XDF1	A	AC	텀	텀	덥	띰	Z S	SO	႘	8	ŏ	Z.	æ	2	ž	₽	₽ B	æ	RA	Æ	Ş	돖	Z.	Z.	КЪ	ည	ž	ΚĀ	RA	RA	RA	RI	RI	7 0	ž ä	DR

```
08X7N5
                                                                                                                                               241
                                                                     121
                                                                                                           181
                                  61
                                                                                                                                                                                                                                 RESULT
                                                                                      g
                                                                                                                                                               g
                                                                                                                                                                                                   g
                                  ö
                                                 g
                                                                     ò
                                                                                                          ð
                                                                                                                          g
                                                                                                                                              ð
                                                                                                                                                                                  à
                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         of the OmpC
                                                                                                                                                                                       Pascual A.,
                                                                                                                                                    WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                             TNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT 300
                                                                                                                                                                                                                                                                           Gaps
                                                                                                                  9
                                                                                                                                   82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                         EYEGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI
                                                                                                                1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Enterobacter.
                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             76.9%; Score 1390.5; DB 2; Length 351; llarity 77.4%; Pred. No. 6.7e-78; Conservative 23; Mismatches 43; Indels 11;
                                                                             Length 362;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ballesteros F., Domenech-Sanchez A., Martinez-Martinez L., I Conejo M.C., Benedi V.J.;
"Genetic characterization and role in ss-lactam resistance and omps Enterobacter cloacae porins.";
Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AJ316540; CAG48383.1;
InterPro; IPR001702; Porin_gram-ve.
Pram; PR00267; Gram-ve.porins; 1.
Probom; PD000808; OMP-2. 1.
PROSTIE; PS00576; GRAM_NEG_PORIN; UNKNOWN_1.
                                                                                               Indels
                                                          9E6ACF4B9DAA8214 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 22 POTENTIAL.
351 AA; 38465 MW; 045617CB302F8968 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                       Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                      YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                 323 YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 362
                                                                             Score 1789; DB 16;
Pred. No. 2.5e-102;
                                                                                                                                                                                                                                                                                                                                                                      351 AA
                                                                                      Pred. No. 2.56
0; Mismatches
InterPro; IPR001702; Porin_gram-ve. Pfam; PF00267; Gram-ve_porins; 1. PRINTS; PR00182; ECOLNEIPORIN. ProDom; PD000808; OMP_2; 1. PROSITE; PS00576; GRAM_NEG_PORIN; 1. Complete proteome. SEQUENCE 362 AA; 39361 MW; 9E6ACF4
                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                            98.98;
99.48;
                                                                                                                                                                                                                                                                                                                                                                                       01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-JUN-2002 (TrEMBLrel. 21,
                                                                                   Best Local Similarity 99.4
Matches 338; Conservative
                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                      (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                  OmpF porin precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                     Enterobacter cloacae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-ATCC13047;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=550;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Porin; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 263;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                    093K98
093K98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                               203
                                                                                                                                                                                                                                                                  241
                                                                                                                                                    19
                                                                                                                                                                                                                                                                                                                                                                                                                           OMPF
                                                                                                                                                                                                                                                                                                                                                   RESULT
Q93K98
  DR WX XX
                                                                                                                 ò
                                                                                                                                 g
                                                                                                                                                    à
                                                                                                                                                                    a
                                                                                                                                                                                        à
                                                                                                                                                                                                         임
                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                  요
```

```
STRAIN-0157:H7 / RIMD 0509952;
XX MEDLINE-21156231; PubMed-11258796;
A Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishlii K., Yokoyama K.,
A Hayashi T., Makino E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Ida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
A Lida T., Takami H., Honda T., Sasakawa H.;
Complete genome sequence of enterchemorrhagic Escherichia coli
T. O15:H7 and genomic comparison with a laboratory strain K-12.";
DNA Res. 8:11-22(2001).
BEBL: AR005202; AAG54566.1;
DNA Res. 8:11-22(2001).
RILEPTO: IPR001702; POTIL_Gram-ve.
PEBL: AR005751; BAB33601.1;
PRINTS: PR00182; ECCLNEIPORIN.
PEMR: PRODORS: CCLNEIPORIN:
PRODOM: PRODORS: CCLNEIPORIN:
PRODOM: PODORS: CRAM_NEG_PORIN: 1.
                                                                                                                                                                                                DTAYSDDFFVGRVGCVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISY 180
                                                                                                                                                                                                                                                              197
                                                                                                                                                                                                                                                                                                                    240
                                                                                                                                                                                                                                                                                                                                                                                                                           TNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEVGAT 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                         258 DN-----GFANKTQDFSVVAQYQFDFGLRPSIAYYKSKAKDVEGIGDEDYINYIDIGAT 311
                                                                                                                257
                                77
23 AEIYNKDGNKLDLYGKAVGLHYFSDND----GNDGDKTYARLGFKGETKINDQLTGYGQ
                                                                                                                                                                                                                                                                                                              EYEGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNATPI
                                                                                                                                                                                                                                                                                                                                           "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."; Nature 409:529-533(2001).
                                                                                  WEYNFOGNNSEGADAOTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG
                                                                                                                                                                                                                              SEQUENCE FROM N.A.
STRAIN-0157:H7 / EDL933 / ArCC 700927;
MEDLINE-21074935; PubMed-11206551;
Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
E (E,Ic,NmpAB) (Outer membrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38966 MW; 3E58F4F284B5D42E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      301 YYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Out-JUN-2002 (TrEMBLrel. 21, Outer membrane pore protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAR-2002 (TrEMBLrel. 20, 01-MAR-2002 (TrEMBLrel. 20, 01-JUN-2002 (TrEMBLrel. 21,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protein PhoE).
PHOE OR 20302 OR ECS0268.
Escherichia coli 0157:H7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              351 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=83334;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Complete proteome
SEQUENCE 351 AA
```

ô

Length 351;

DB 16;

Score 1113;

61.5%;

Query Match

ന

ò g õ å

ò

ò

```
Conejo M.C., Benedi V.J.;
"Genetic Characterization and role in ss-lactam resistance of the Ompc and ompt Enterobacter cloadea porins";
Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AJ316539; CAC48382.1:
InterPro; IPR001229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
InterPro; IPR00157; Gram-ve_porins; 1.
Probom; PD000808; OMP_2: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pascual A.,
                      DT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKN------ERDTARRSN 170
                                                                                                                    195 GDGFGLSSSYEIGAGVSVGAAYASSNRTLAQKNSTFGKGDKADAWTTGLKYSDNGVYLAA 254
                                                                                                                                                                                                                                                                                                                 6.1 WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12.1 DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGK------NERDTARRSN 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22 AEIXNKDGNKLDLYGKVDGLHYFSDDDSQ-----DGDQTYMRLGFKGETQVNDQLTGYGG 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                  284 ECIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                        171 GDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAA
                                                                                                                                                                                                                                                                                      NYGETRNATPIT - - - - - NKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        17.1 GDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; \operatorname{Proteobacteria}; \operatorname{gamma} \operatorname{subdivision}; \operatorname{Enterobacteriaceae}; \operatorname{Enterobacter}.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 363;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-ATCC13047;
Ballesteros F., Domenech-Sanchez A., Martinez-Martinez L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  77; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL, 681370F63C3075B6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 58.7%; Score 1061; DB 2; Best Local Similarity 60.6%; Pred. No. 1.1e-57; Matches 215; Conservative 35; Mismatches 77;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GRAM_NEG_PORIN; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROSITE; PS00576; GRAM_NEG_PORIN; UNKNO
PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21 FO
39940 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19,
19,
21,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-DEC-2001 (TrEMBLrel.
01-DEC-2001 (TrEMBLrel.
01-JUN-2002 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OmpC porin precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Enterobacter cloacae,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    363 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBL_TaxID=550;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Porin; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q93K99
Q93K99;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNAL
                                                                                                                                                                                                                                                                                                                                  255
                                                                                                                                        136
                                                                                                                                                                                                                                                                                      230
                                                                                                                                   셤
                                                                                                                                                                                       ò
                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCOORDING TO THE PROPERTY OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                               5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-CO-92 / BIOVAR ORIENTALIS;

MEDLINE=21470413; PubMed=11586360;

Parkhill J., Wren B.W., Thomson N.P., Churcher C., Mungall K.L.,

A Prentice M.B., Sebaiha M., James K.D., Churcher C., Mungall K.L.,

Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,

Chillingworth T., Crouin A., Davise R.M., Davise P., Dougan G.,

Reltwell T., Hamin N., Holroyd S., Jagels K., Karlyshev A.V.,

Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,

Simmonds M., Skelfron J., Stevens K., Whitehead S., Barrell B.G.,

Mature 413:523-527(2001).

Rembi: AJ41447; CASC90043.1; -.

InterPro; IPR001702; Porin_gram-ve.

Rembi: AJ41447; CASC90043.1; -.

Rempi: PR001702; Profin_gram-ve.

Rempi: PR00182; ECCLNEIPORIN.

Render S. PR00182; ECCLNEIPORIN.

Respons PR000888 OMP_2: 1.

Respons PR000888 OMP_2: 1.

Rempi: PR000888 OMP_2: 1.
                                                                                                                                                                                     61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                121 D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                            238 TPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                               180 YEYEG--FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                    22 AEIYNKDGNKLDVYGKVKAMHYMSDNDSK-----DGDQSYIRFGFKGETQINDQLTGYGR 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22 AEIYNKDGNKLDLYGKVKALHYFS----DNTKSDGDKSYVRLGFKGVTQITDELSGYGQ 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                     1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 16; Length 371;
Pred. No. 7e-61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40708 MW; 1AEB9F1E343263D2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             309 GATYYFNKNMSAFVDYKINQLDSDNKLNINNDDIVAVGMTYQF 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score Luoz,
Pred. No. 6e-59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                48; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              59.8%; Score 1082; 59.4%; Pred. No. 6e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                          38;
                   62.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAR-2002 (TrEMBLrel. 20, 01-MAR-2002 (TrEMBLrel. 20, 01-JUN-2002 (TrEMBLrel. 21,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Outer membrane protein C2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                212; Conservative
                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
              Best Local Similarity
Matches 213; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    371 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pestis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=632;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Yersinia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Yersinia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           985280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

;

ð

ð g å a

```
311 GGADNPAGVDDKDLVKYADVGATYYENKNMSTYVDYKINLLDEDDSFYTANGISTDDIVA 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           168 RSNGDGVGGSISYEY-EGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIY 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 DT-AYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNE--------RDTAR 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         227 LAANYGETRNATPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVE-- 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVA 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 AEVYNKDGNKLDLYGKVDGLHYFSDNSAK----DGDQSYARLGFKGETQINDQLTGYGG 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Prilipov A., Phale P.S., Koebnik R., Widmer C., Rosenbusch J.P.; Identification and characterization of two quiescent porin genes, nmpC and ompN, in Escherichia coli BE."; J. Bacteriol. 180:3388-3392(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
NCBI_TaxID-562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             57.3%; Score 1036; DB 2; Length 377; 57.2%; Pred. No. 4.1e-56; tive 41; Mismatches 78; Indels 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL, AF035618; AAC19644.1; -...
HSSP; P02931; 1GFN.
InterPro; IPR001329; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
InterPro; IPR001703; Reg_chr_condens.
Pfam; PF00267; Gram-ve_porins; 1.
PRINTS; PR001808; OMP_2; 1.
PROSITE; PS00626; RCC1_2; UNKNOWN_1.
SEQUENCE 377 AA; 41236 MW; 7F4D681A2BC10F8C CRC64;
                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                        377 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         372 AA
                                                                                                                                                                                                                                                                Created)
                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
STRAIN-BE;
MEDLINE-98317278; Pubmed-9642192;
                                                                                                                                                                                                                                                           (TrEMBLrel. 0 (TrEMBLrel. 0 (TrEMBLrel. 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                        Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      334 VGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LGLVYQF 377
     335 GIVYOF 340
                                                           369 GLVYOF 374
                                                                                                                                                                                                                                                           01-NOV-1998 (
01-NOV-1998 (
01-JUN-2002 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 210;
                                                                                                                                                                                                                                                                                                                                                      Porin OmpN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q93K39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
Q93K39
ID Q93K3
                                                                                                                                             RESULT 7
085030
                                                              셤
        ö
                                                                                                                                                                                                                                      SO DE REPRESENTATION OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A EDIANGE-99235749; PubMed-10217760;

RA MEDIALE-99235749; PubMed-10217760;

RA Benedi V.J., Alberti S.,

Ra Benedi V.J., Alberti S.,

Ra Genedi V.J., Alberti S.,

Ra Tidentification and characterization of a new porin gene of Klebsiella RT The mountae: Its role in beta-lactam antibiotic resistance.";

J. Bacteriol. 181:2726-2732(1999).

Ra Beneli A.J011502: RAA09666.1; -.

R HSSP; Q48473; 10SM.

R HSSP; Q48473; 10SM.

R HSSP; Q48473; 10SM.

R InterPro; IPR001702; Porin, gram-ve.

R InterPro; IPR001702; Porin, gram-ve.

R InterPro; IPR001803; Decining: 1.

R PRINTS; PR00182; ECOLNEIPORIN.

R PRODOM; P0000808; OMP_2; 1.

R PODOM; PS00626; RCC1_2; UNKNOWN_1.

R POTIN; Signal.

SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6
230 NYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVE-GIGD 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---RRSNGDGVGGSISYEYEGFGI--VGAYGAADRINLQEAQPLGNGKKAEQWATGLKYD 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ANNIYLAANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAK 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 135 DSYTYADNFMAGRANGVATYRNSDFFGLVEGLNFALQYQGKNEGQNAQDINVGTNNRSSD 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DVEGIG---DVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAV 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 ABIYNKDGNKLDLYGKVDGLHYFSSDSKK----DGDQTYLRFGFKGETQINDILLGYGQ 76
                                   289 VDLVNYFEVGATYYFNKNMSTYVDYIINQIDS---DNKLGVGSDDTVAVGIVYQF 340
                                                                                                                                             Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 2; Length 374;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   57.7%; Score 1044.5; DB 2; Length 3
58.5%; Pred. No. 1.2e-56;
Live 37; Mismatches 76; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE 374 AA; 41612 MW; B66DF0D74A1A169B CRC64;
                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OMPK37 PORIN.
                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 58.58
Matches 214; Conservative
                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21
374
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ompk37 porin precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Klebsiella pneumoniae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=573;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Klebsiella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                             087754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      309
                                                                                                                                                                                                                                                        RESULT 6
```

A PART HAND DAY OF THE PART OF

ò g g

ò

g

ð

q

ò

g

ò

õ

9

Gaps

Thu

```
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDT-------ARR 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AANYGETRNATPIT -- NKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEG 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SNGDGVGGSISYE-YEGFGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYL 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 AEIYNKDGNKLDLYGKIDGLHYFS-----DDKSVDGDQTYMRVGVKGETQINDQLTGYGQ 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           286 IGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDN---KLGVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C Ompc).
                                                                                    Klebsiella pneumoniae.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                              25;
                                                                                                                                                                                                                                                                                                                                                                                                                 DB 2; Length 372;
                                                                                                                                                                                     Crowley B., Benedl V.J., Domenech-Sanchez A.;
"Porin deficiency and SHV-2 expression results in increased re
to cephalosporins and carbapenems in Klebsiella pneumoniae.";
submitted (AUG-201) to the EMBL/GenBank/DDBJ databases.
EMBL, AJ344089; CAC50885.1;
InterPro; IPR001702; Porin_gram.ve.
InterPro; IPR001702; Porin_gram.ve.
Probom; PD000808; OMP_2; 1.
PROSTIE; PS00576; GRAM_NEG_PORIN; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                              57.2%; Score 1035.5; DB 2; Length 58.4%; Pred. No. 4.3e-56; tive 38; Mismatches 86; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAR-2002 (TTEMBLEEL. 20, Created)
01-MAR-2002 (TTEMBLEEL. 20, Last sequence update)
01-MAR-2002 (TTEMBLEEL. 21, Last sequence update)
01-JUN-2002 (TTEMBLEEL. 21, Last annotation update)
Outer membrane protein lb (Ib,c) (Outer membrane protein
OMPC OR 23473 OR ECS3104.
Escherichia coll 0157:H7.
                                                                                                                                                                                                                                                                                                                                                                    22 372 OMPK36.
372 Aa; 40753 MW; 94B84DA54AF97736 CRC64;
            01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                367 AA
                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
STRAIN-0157:H7 / EDL933 / ATCC 700927;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 209; Conservative
                                                        OmpK36 protein precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                     21
372
                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=83334;
                                                                                                                Klebsiella.
NCBI_TaxID=573;
                                                                                                                                                                             STRAIN-103624;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Escherichia
                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                          Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                  Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q8XE41;
                                                                                                                                                                                                                                                                                                                                                        SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q8XE41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08XE41
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
CSTRAIN-O15:H7 / RIMD 0508952;

X MEDLINE-21156231; PubMed-11258796;

X Hayash1 T. Makino K., Ohinishi M., Kurokawa K., Ishii K., Yokoyama K., Hayash1 T., Makino K., Ohinishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Ishara S., Siba T., Hattori M., Shinagawa H.;

X Kuhara S., Siba T., Hattori M., Shinagawa H.;

X Kuhara S., Siba T., Hattori M., Shinagawa H.;

X Kuhara S., Siba T., Hattori M., Shinagawa H.;

X Romplete genome sequence of enterohemorrhagic Escherichia coli

O15:H7 and genomic comparison with a laboratory strain K-12.";

NR EMBL; AR0035453 AAG57350.1; -.

REMBL; AR0035453 AAG57350.1; -.

REMBL; AR001209; OMP_2.

InterPro; IPR001202; OMP_2.

InterPro; IPR001202; OmP_2.

RIMERPRO; PR00182; OmP_2.

REMBL; AR00182; Cama-ve_porins; 1.

PROSITE; PS000576; GRAM_NEG_CORIN; 1.

PROSITE; PS000576; GRAM_NEG_CORIN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WEYNFOGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGK------NERDTARRSN 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22 AEVYNKDGNKLDLYGKVDGLHYFS----DDKSVDGDQTYMRLGFKGETQVTDQLIGYGG 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
MEDLINE-21074935; PubMed-11206551;
Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Posse D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posten G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K., Anatharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;
"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.",
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GDGVGGSISYEYEGFGIVGAYGAADRTNLQEAQPL--GNGKKAEQWATGLKYDANNIYLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----DVDLVNYFEVGATYYFNKNMSTXVDYIINQIDSDNKL----GVGSDDTVAVGIVYQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 367;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               367 AA; 40508 MW; 028A0146D462CAA4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              57.1%; Score 1033; DB 58.7%; Pred. No. 6e-56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                43; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (TrEMBLrel. 15, C
(TrEMBLrel. 15, I
(TrEMBLrel. 21, I
                                                                                                                                        Nature 409:529-533(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Complete proteome. SEQUENCE 367 AA;
                                                                                                                                                                         SECUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-OCT-2000
01-OCT-2000
01-JUN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Sim
Matches 212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          340 F 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09K597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9K597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09K597
        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               44446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

ő

```
SEQUENCE FROM N.A.
                                    LOOS M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9RH85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09RH85
                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                              134
                                                                                                                                                                                                                                                                                                                                                                                                                                                          228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
    셤
                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .;
8
                                                                                                                      Submitted (AUG-2000) to the EMBL/GibbJ databases.

-1- SUBMITT: HOMOTRIMER (BY SIMILARITY).

-1- SUBCLIULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE

(BY SIMILARITY).

(BY SIMILARITY).

RHSSP: OS2503: IIIV.

RICEPPO: IPR001229; OMP_2.

RICEPPO: IPR001702; Porin_gram.ve.

RICEPPO: IPR001702; Porin_gram.ve.

RICEPPO: IPR001702; Porin_gram.ve.

RICEPPO: IPR00180; Camu-ve_porins; 1.

RPRINTS; PR00182; ECOLMEIPORIN; 1.

RPRODOM: PR00182; ECOLMEIPORIN; 1.

RR PROSTIE; PS00030; RRM_RNP_1; UNKNOWN_1.

Outer membrane; Porin; Signal; Transmembrane.

T. SIGNAL.

T. SIGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                    DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGK------NERDTARRS 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                        NGDGVGGSISYEYEGFGIVGAYGAADRT-NLQEAQPLGNGKKAEQWATGLKYDANNIYLA 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 254 AQYTQTYNATRV----GSLGWANKAQNFEAVAQYQFDFGLRPSLAYLQSKGKNLGRGYD 308
                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ANYGETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV-EGIG 287
                                                                                                                                                                                                                                                                                                                                                                       1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Low A.S., MacKenzie F.M., Gould I.M., Booth I.R.;
"Parallel evolution of multi-resistant bacteria in a patient with recurrent septicaemia: unique data that support the presence of separate protected environments.";
                              Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Salmonella enterica subsp. enterica serovar Minnesota.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
NCBI_TaxID=70803;
                                                                                                                                                                                                                                                                                                                                         31;
                                                                                                                                                                                                                                                                                                                    DB 2; Length 364;
                                                                                                                                                                                                                                                                                                                                         77; Indels
                                                                                                                                                                                                                                                                                               9B583F2C11344E31 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-UTN-2002 (TrEMBLrel. 21, Last annotation update)
00. Outer membrane protein C precursor.
                                                                                                                                                                                                                                                                                                                  56.9%; Score 1029.5; DB 58.5%; Pred. No. 9.8e-56; ive 40; Mismatches 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 378 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
 membrane porin C precursor
                                                                                                                                                                                                                                                                                               364 AA; 40312 MW;
                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 58.5%
Matches 209; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                   Escherichia coli
                                                 NCBI_TaxID-562;
                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09K3E6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09K3E6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 11
Q9K3E6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               229
ð
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
```

```
WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            173 GVGGSISYEY-EGFGIVGAY----GAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYL 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E----GIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDN---KLGVGSDDTVAVG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    314 SNGYGASYGDQDIVKYVDVGATYYFNKNMSTYVDYKINLLDKNDFTRDAGINTDDIVALG 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKN-----ERDTAR---RSNGD 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22 AEIYNKDGNKLDLFGKVDGLHYFS----DDKGSDGDQTYMRIGFKGETQVNDQLTGYGQ 76
                                                              -1- SUBBINIT: HOMOTRIMER (BY SIMILARITY).
-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AANYGETRNATPI----TNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OMPC.
Escherichia coli 0157:H7.
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2; Length 378;
Zimmermann H., Wassenaar T.M., Laubenheimer-Preusse H., Petry
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      56.9%; Score 1029.5; DB 2; Length 57.0%; Pred. No. 1e-55; Attive 46; Mismatches 78; Indels
                                             Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                          SIGNAL 1 21 POTENTIAL.
CHAIN 22 378 OUTER MEMBRANE PROTEIN C.
SEQUENCE 378 AA; 41269 MW; 254524EB9EC3849C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                   Outer membrane; Porin; Signal; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   366 AA
                                                                                                                                                                                  Interpro; IPR001229; OMP_2.
Interpro; IPR001702; Porin_gram-ve.
Interpro; IPR001504; RNA_rec_mot.
Pfam; PF00267; Gram-ve_porins; 1.
PRINTS; PR00182; ECCLNEIPORIN.
PRODOM; PD0008086; OMP_2, 1.
PROSITE; PS00576; GRAM_NEG_PORIN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                       EMBL; Y15844; CAB96613.1; -. HSSP; 052503; 1IIV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Outer membrane protein OmpC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 208; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=0157:H7;
Yu S.L., Syu W.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=83334;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : | | | | | 374 LVYQF 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          336 IVYOF 340
```

```
à
                                                                                                                                                                                셤
                                                                                                                                                                                                        ç
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                   DTAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGK------NERDTARRSNG 171
                                                                                                                                                                                                                                                                                                                                                                172 DGVGGSISYEYEGFGIVGAYGAADRINLQEAQPL--GNGKKAEQWATGLKYDANNIYLAA 229
                                                                                                                                                                                                                                                                                                                                                                                                                           252
                                                                                                                                                                                                                                                                                                                                                                                                                                                230 NYGETRNATPITNKFTNTSGFANKTODVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIG-- 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                            253 QYTQTYNATRY----GSLGWANKAQNFEAVAQYQFDFGLRPSLAYLQSKGKNLGVINGR 307
                                                                                                                                                                                                                                                                  22 AEVYNKDGNKLDLYGKVDGLHYFS-----DDKSVDGDQTYMRLGFKGETQVTDQLTGYGQ 76
                                                                                                                                                                                                                                Gaps
*Altered outer membrane protein OmpC in hemorrhagic Escherichia coli
                      Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Phylogenetic analysis of Calymmatobacterium granulomatis.";
Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
-!- SUBURIT: HOMOTRIMER (BY SIMILARITY).
-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          308 NYDDEDILKYVDVGATYYFNKNMSTYVDYKINLLD-DNQFTRDAGINTDNIVALGEFYQF
                                                                                                                                                                                                                                                     1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGCNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                               --- DVDLVNYFEVGATYFNKNMSTYVDYIINQIDSDNKL----GVGSDDTVAVGIVYQF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Klebsiella pneumoniae (subsp. ozaenae).
Bacteria: Proteobacteria: gamma subdivision; Enterobacteriaceae;
Klebsiella.
                                                                                                                                                                                                                             35;
                                                                                                                                                                                                      Length 366;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Carter J.S., Bowden F.J., Bastian I., Myers G.M., Sriprakash
Kemp D.J.;
                                                                                                                                                                                                                             71; Indels
                                                                                                                                                                 Outer membrane; Porin; Signal; Transmembrane.
SEQUENCE 366 AA: 40499 MW; 6A4EAD1652565C00 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                    DB 2;
                                                                                                                                                                                                    56.9%; Score 1028.5; DB 58.6%; Pred. No. 1.1e-55; live 43; Mismatches 71
                                                       (BY SIMILARITY).

EMBL; AF057355; AAF21761.1; -.

HSSP, OS2503; 111V.

INTERPRO; IPR001329; OMP_2.

INTERPRO; IPR001702; Porin_gram-ve.

Pfam; PF00267; Gram-ve_porins; I.

PR.NYTS; PR00182; Gram-ve_porins; I.

PRODOM; PD000808; OMP_2; I.

PROSITE; PS00576; GRAM-NEG_PORIN; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AF009229; AAD21517.1; -.
HSSP; P09323; IPD.
InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Phosphate porin (Fragment).
                                                                                                                                                                                                                             Matches 211; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=574;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-1436;
                       Submitted
                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                   Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9RAW3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 O9RAW3
                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
Q9RAW3
유
                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121 D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                       15 AEVYNKNANKLDVYGKIKAMHYFSDYDSK----DGDQTYVRFGIKGETQINDDLTGYGR 69
                                                                                                                                                                                                                                                                                                                             1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Phylogenetic evidence for reclassification of Calymmatobacterium granulomatis as Klebsiella granulomatis comb. nov.";
Int. J. Syst. Bacteriol. 49:1695-170(1999).
-1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Carter J.S., Bowden F.J., Bastlan I., Myers G.M., Sriprakash K.S.,
Kemp D.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180 YEYEG--FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             238 TPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Klebsiella pneumoniae subsp. rhinoscleromatis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                              DB 2; Length 333;
                                                                                                                                                                                                                            56.5%; Score 1022.5; DB 2; Length
58.9%; Pred. No. 2.4e-55;
Live 41; Mismatches 79; Indels
                                                                                                                                                                              36839 MW; 68AF347207770ABA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AF009230; AAD21518.1; -.
HSSP; P02932; IPH0.
InterPro; IPR001229; ONP_2.
InterPro; IPR001209; Porin_gram-ve.
Pfam; PF00267; Gram-ve_porins; 1.
ProDom; PR000808; OMP_2: 1.
PROSTE; PS00576; GRAM.NEG_PORIN; 1.
PROSTE; PS00576; GRAM.NEG_PORIN; 1.
Pfam; PF00267; Gram-ve_porins; 1.
PRINTS; PR0182; ECOLNEIPORIN.
ProDom; P0000808; OMP_2; 1.
Outer membrane; Porin; Signal; Transmembrane.
NON_TER 33
EQUENCE 333 AA; 36839 MW; 68AF347207770AI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        333 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDD 330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-NCTC5046;
MEDLINE-20023050; Pubmed-10555350;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Phosphate porin (Fragment).
                                                                                                                                                                                                                              Query Match 56.5%
Best Local Similarity 58.9%
Matches 196; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-39831;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Klebsiella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09RAW2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     O9RAW2
```

œ

```
ò
                                                                                                                                 g
                                                                                                                                                                                   ò
                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                       .;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8;
                                                                                                                                                                                                                                                                                                                             D-TAYSDDFFVGRVGGVATYRNSNFFGLVDGLNFAVQYLGKNERDTARRSNCDGVGGSIS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                61 WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                            WEYNFQGNNSEGADAQTGNKTRLAFAGLKYADVGSFDYGRNYGVVYDALGYTDMLPEFGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YEYEG--FGIVGAYGAADRINLQEAQPLGNGKKAEQWATGLKYDANNIYLAANYGETRNA 237
                                                                                                                                       Gaps
                                                                                                                                                                                                                 15 AEVYNKNANKLDVYGKIKAMHYFSDYDSK-----DGDQTYVRFGIKGETQINDDLJGYGR 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                      1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23 AEIYNKDGNKLDLYGKAVGRHVWTT-TGDSK---NADQTYAQIGFKGETQINTDLIGFGQ 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 AEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFKGETQINSDLTGYGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TPITNKFINTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSKAKDVEGIGDVDLVNYFEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Petry F.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Salmonella enterica subsp. enterica serovar Minnesota.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
                                                                                                                                    17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13;
                                                                                DB 2; Length 333;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56.5%; Score 1022.5; DB 2; Length 363; 57.3%; Pred. No. 2.6e-55; Live 54; Mismatches 81; Indels 13;
                                                                             56.5%; Score 1022.5; DB 2; Length 58.9%; Pred. No. 2.4e-55; Live 41; Mismatches 79; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIRMATMESFILLI;
Zimmermann H., Wassenaar T.M., Laubenheimer-Preusse H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
EMBL; Y15843; CAB96614.1; -.
HSSP; PO2931; IGFN.
InterPro; IPR003229; OMP_2.
InterPro; IPR001702; Porin_gram-ve.
Pfam; PF00267; Gram-ve_porins; 1.
PRINTS; PR00182; ECOLNEIPORIN.
PRODOM: PD000808; OMP_2; 1.
PROSIȚE; PS00576; GRAM_NEG_PORIN; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 22 POTENTIAL.
23 363 OUTER MEMBRANE PROTEIN FX.
363 AA; 40090 MW; F5058DEDEA516859 CRC64;
                        68AF347207770ABA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09K3E7,
01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                363 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      298 GATYYFNKNMSTYVDYIINQIDSDNKLGVGSDD 330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Outer membrane protein Fx precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
333
36839 MW;
                                                                          Query Match
Best Local Similarity 58.9%
Matches 196; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 57.39
Matches 199; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
333 33
333 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=70803;
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09K3E7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
Q9K3E7
                                                                                                                                                                                                                                                                                                                                                                                                     121
FT
                                                                                                                                                                                      ò
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCOCCOORDITATION OF THE READ BRANCH 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
```

-

```
Peptide used to as Synthetic self-ass 21p5 polypeptide use combinant ubideul Human procaspase B Human caspase-B. Recombinant ubidul Spider recombinant Spider recombinant Recombinant Spider secombinant Human procalcitoni Human procalcitoni Human procalcitoni Human apoAIV mutei
                                                                                                                                                                                                                                                                                                                              P. aeruginosa H co
P. aeruginosa H co
P. aeruginosa H co
H6FXTTIBB-UB fusio
Collagen binding p
P. falciparum spor
P.aeruginosa E col
Collod coll protei
Sequence of an epi
                           Peptide used to as HIV protease-assoc WIV protease-assoc Synthetic Self-ass Amino acid sequenc Amino acid sequenc Amino acid sequenc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Recognin Bl; structural protein; purification; fibre; spinning.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ALIGNMENTS
                                                                                                                                                                                                                                                                                                     AABB6B28
AAR39486
AAY44384
AAY44380
AAY44382
                                                                                                                                                                                                                                                                                                                                                                       AAW94266
AAW31552
AAR13992
AAY44385
AAY44381
AAB10560
AAB10560
                                                                                                                                                                             ABB83407
AAY56997
                                                                                                                                                                                                                                                                           AAY44383
AAB86825
                                                                                                                                                                                                                                    AAB82608
                                                                                                                                                                                                                                                AAB82609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB82613 standard; Protein; 74 AA
 32-OCT-2001 (first entry)
 Recognin Bl protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200153333-A1.
124
121
121
119.5
118.5
118.5
114.5
113
111.1
104
103
                                                                                                                                                                                         101.5
101.5
101
97
97
                                                                                                                                                                                                                                                           96.5
                                                                                                                                                                                                                                                                                     95.5
95.5
95
                                                                                                                                                                                                                                                                                                                                                                                                                 91.5
91.5
91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                              \frac{9}{4}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB82613;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 1
AAB82613
ID AAB8
 December 18, 2002, 16:05:10 ; Search time 4.94091 Seconds (without alignments) 1995.696 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                           A_Geneseq_101002:*

| SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDS2/gcgdata/geneseqg-embl/AA1981.DAT:*
| SIDS2/gcgdata/geneseqg-embl/AA1991.DAT:*
| SIDS2/gcgdata/geneseqg-embl/AA1991.DAT:*
| SIDS2/gcgdata/geneseqg-embl/AA1991.DAT:*
| SIDS2/gcgdata/geneseqg-embl/AA1991.DAT:*
| SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1991.DAT:*
| SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1991.DAT:*
| SIDS2/gcgdata/geneseqg/geneseqg-embl/AA1991.DAT:*
| SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1995.DAT:*
| SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1996.DAT:*
| SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1996.DAT:*
| SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1996.DAT:*
                                                                                                                                                                        1 MRGSHHHHHHGSMASGDLKN.....AKIGDLNNTSGIRRPAAKLN
            GenCore version 5.1.3 Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                       Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                             908470 seqs, 133250620 residues
                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                 protein search, using sw model
                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                  length: 0
length: 2000000000
                                                                                                                                          US-09-490-291-11
377
```

sed

Minimum DB s Maximum DB s

Database

Title: Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Recovering structural polypeptides in a biological sample, useful for

Butler

Wello CM, Arcidiacono S,

Recognin Bl protei Copolymer block pr Copolymer block pr Artificial recogni Artificial recogni Artificial recogni Colled coil protei Artificial recogni Artificial recogni Artificial recogni

AAB10562 AAB10561 AAW43025 AAW43026 AAW43027 AAB10558 AAW43024 AAW43024 AAW43024

74.1 64.6 58.6 56.5 52.8 52.3 47.0 48.0

377 279.5 243.5 221 221 203 1099 1197 1181

10 9 8 7 9 9

AAB82613

DB

Query Match Length

Score

Description

SUMMARIES

WPI; 2001-483136/52. N-PSDB; AAH26305.

(MELL/) MELLO C M.
(ARCI/) ARCIDIACONO S.
(BUTL/) BUTLER M M.
(USSA ) US SEC OF ARMY.

01-NOV-2000; 2000WO-US30086 20-JAN-2000; 2000US-0490291

26-JUL-2001.

```
The present sequence is that of synthetic coiled protein Recognin B1. The invention provides methods for purifying structural proteins, including spider slik proteins. Organic acids are used to lyse recombinant cells or other biological samples (such as non-recombinantly derived cells), and enrich the purity and yleids of structural proteins by hydrolysing many of the macromolecules of structural proteins intact. In the present case, while leaving the structural proteins intact. In the present case, Recognin B1 DNA was cloned into vector pQB-9 for recombinant protein was purified from lyophilised pellets using formic acid, acetic acid, propionic acid, butyric acid or valeric acid. Formic acid, acetic acid, propionic acid, butyric acid or valeric acid, were able to extract acids, but none of the other organic acids, were able to extract acids was able to extract Recognin B1 in a relatively pure form. The new method has the following advantages over prior art: it involves fewer steps, requires less time and smaller volumes of reagents, ceasits in better recovery of protein at higher purity (70-99%), is being solution reducing hazardous waste accumulation and cost.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Colled coll protein; copolymer blocker; vulnerary; wound treatment; integrin; fibroblast cell; matrix formation; skin regeneration; scaffold; antibiotic; dressing; abrasion; burn.
purifying and spinning spider silks and other structural proteins, comprises treating the sample containing the polypeptides with an acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MRGSHHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100.0%; Score 377; DB 22;
100.0%; Pred. No. 1e-32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Copolymer block protein insert L2-B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Tirrell DA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB10562 standard; protein; 76 AA.
                                                                     2; Page 44; 49pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97US-0956307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97US-0956307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (UYMA-) UNIV MASSACHUSETTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-DEC-2000 (flrst entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Petka WA, McGrath KP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 NNTSGIRRPAAKLN 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                  74 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unidentified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US6090911-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB10562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                     Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

useful for

Synthetic block copolymer with an antibiotic compound, usefudressing abrasion, burn or non-puncture wound, comprises two

WPI; 2000-542235/49.

ö

```
This invention describes a novel synthetic block copolymer XYZ comprising two alpha-helical protein blocks X and Z, a water soluble, random-coil two alpha-helical protein blocks X and Z, and inker proteins for linking the alpha-helical protein blocks to the random-coil protein block. The products of the invention have vulnerary activity. A copolymer produced from a gene that encodes a derivative of AC_1OA copolymer produced in water to form a gel, which was represed to treat a wound. The cell binding domain was the integrin ArgGlyAspSer sequence which is known to bind to gpila/IIIa proteins expressed on fibroblast secone centrapped within the gel and thus remain at the site of the wound, a synthetic block copolymer together with an antibiotic compound is useful for wound dressing. The copolymer is useful for dressing copolymer together with an antibiotic compound is useful for wound dressing. The copolymer is useful for dressing bolymers of equal size are produced biologically from a single template. The intermolecular binding of alpha helical blocks are monodisperse which gives a uniform pore size depending on the length of random-coil block. The new copolymer block protein L2-A which is used to illustrate the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Coiled coil protein; copolymer blocker; vulnerary; wound treatment; integrin; fibroblast cell; matrix formation; skin regeneration; scaffold; antibiotic; dressing; abrasion; burn.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MRGSHHHHHHGS-----MASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              alpha-helical protein blocks, random-coil protein block and linker
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 21; Length 76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 279.5; DB 2:
Pred. No. 2.2e-22;
1; Mismatches 1.
                                                  Example 1; Column 33-36; 53pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Copolymer block protein insert L2-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Petka WA, McGrath KP, Tirrell DA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB10561 standard; protein; 76 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        74.18;
89.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-0956307.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (UYMA-) UNIV MASSACHUSETTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 89.1
Matches 57; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                     method of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-542235/49.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     76 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               56 KIGD 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ||||
61 KIGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US6090911-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB10561;
                  proteins
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
```

```
This invention describes a novel synthetic block copolymer XYZ comprising two alpha-helical protein blocks X and Z, and linker proteins for linking the alpha-helical protein blocks to the random-coil protein block. The alpha-helical protein blocks to the random-coil protein block. The products of the invention have vulnerary activity. A copolymer produced from a gene that encodes a derivative of AC_1.0A copolymer produced in water to form a gel, which was then used to treat a wound. The cell binding domain was the integrin ArgolyAapSer sequence which is known to bind to gplia/Tia proteins expressed on fibroblast cells necessary for matrix formation for the regeneration of skin. Fibroblasts become entrapped within the gel and thus remain at the site of the wound, serving as a scaffold for the regeneration of tissue surrounding the wound. A synthetic block copolymer together with an antibiotic compound is useful for wound dressing. The copolymer to useful for dressing abrasion, burn or non-puncture wound. Unlike conventional polymers, the colymers of equal size are produced biologically from a single template. The intermolecular binding of alpha-helical blocks are monodisperse which gives a uniform pore size depending on the length of random-coil block. The new copolymers have low molecular weight hence have decreased viscosity solutions or gels under suitable conditions. This sequence represents a copolymer block protein L2-A which is used to illustrate the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Recognition sequence; self-assembling; protein-based structural material; spontaneous heterodimerisation; coiled-coil alpha helix; selectively permeable membrane; coated fabric.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MRGSHHHHHHGS-----MASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
Synthetic block copolymer with an antibiotic compound, useful for dressing abrasion, burn or non-puncture wound, comprises two alpha-helical protein blocks, random-coil protein block and linker
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .,
.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 21; Length 76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64.6%; Score 243.5; DB 275.0%; Pred. No. 1.4e-18; ive 10; Mismatches 1
                                                                                               Example 1; Column 33-34; 53pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAW43025 standard; Peptide; 84 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Artificial recognition sequence 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    95US-0452592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         93US-0068948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 75.0 nes 48; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        McGrath KP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           76 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             :111
EIGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      56 KIGD 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-MAY-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27-JAN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JS5712366-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kaplan DL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAW43025;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                             proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAW43025
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
```

```
AAW43024-28 represent recognition sequences obtained from the library of recognition sequenes coded for by AAV04943. This DNA sequence is recognition sequenes coded for by AAV04943. This DNA sequence is cobtained by reverse translation. The codons were chosen to maximise expression in Escherichia coil, and to introduce useful restriction for subsequent genetic manipulations. The DNA sequence use a "mixed slite" of approach at the first base of the codons for amino acids positions 5 and 7 of the heptad (see below). The formation of a self-assembling structural material is mediated by the artificial recognition sequences content in sequences. A method of producing a self-assembling protein-based structural material comprises mixing a multiplicity of 2 different artificial peptide sequences, designed so as to spontaneously heterodimerise. Both artificial peptides have a coiled-coil alpha helical secondary structure having at least 2 heptad units. The artificial peptides spontaneously heterodimerials designed for specific functions e.g.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Recognition sequence; self-assembling; protein-based structural material; spontaneous heterodimerisation; colled-coil alpha helix; selectively permeable membrane; coated fabric.
                                                                                                                                                                                                                                                                                                                                                                                                  membranes, fibres, absorbants, reactive materials, etc. A specific application is in the fabrication of selectively permeable membranes for protective gloves and coated fabrics for fuel handlers and other hazardous materials.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hazardous măterials. Materials for medical implants, wound healants and other medical treatments would also be feasible using this technology.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                   Production of peptide hetero:dimer(s) - in the fabrication of self-assembling protein-based structural material(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Production of peptide hetero:dimer(s) - in the fabrication of self-assembling protein-based structural material(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  58.6%; Score 221; DB 19; 97.9%; Pred. No. 3.9e-16; 11ve 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Artificial recognition sequence 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                                        Disclosure; Fig 4B; 17pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAW43026 standard; Peptide; 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                93US-0068948.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95US-0452592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-MAY-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 97.9 Matches 46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          McGrath KP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1998-120029/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            84 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-MAY-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27-JAN-1998.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USS712366-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kaplan DL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAW43026;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAW43026
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

4

 $\overset{\mathsf{A}}{\times} \overset{\mathsf{X}}{\circ} \overset{\mathsf{A}}{\circ} \overset{\mathsf{A}}{\circ$ 

```
expression in Excenting coll, and to introduce user, restriction, coll subsequent genetic manipulations. The DNA sequence use a "mixed site" approach at the first base of the codons for aniho acids positions 5 and coll the heptad (see below). The formation of a self-assembling structural material is mediated by the artificial recognition sequences of structural material comprises mixing a multiplicity of protein-based structural material comprises mixing a multiplicity of 2 different artificial peptide sequences, designed so as to spontaneously heterodimerise. Both artificial peptides spontaneously heterodimerise into a self-assembled coll alpha helical secondary structure having at least 2 heptad units. The artificial peptides spontaneously heterodimerise into a self-assembled protein-based structural material. The method is used to produce consocial structural materials designed for specific functions e.g. consocial structural materials designed for specific functions of protein-based structural materials designed for specific functions is in the fabrication of selectively permeable membranes for protective gloves and coated fabrics for fuel handlers and other coated fabrics for fuel handlers wound healants and other medical treatments would also be feasible using this technology.
                expression in Escherichia coli, and to introduce useful restriction for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 7
             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                  AAW43024-28 represent recognition sequences obtained from the library of recognition sequenes coded for by AAV04943. This DNA sequence is obtained by reverse translation. The codons were chosen to maximise expression in Escherichia coil, and to introduce useful restriction for subsequent genetic manipulations. The DNA sequence use a "mixed site" of approach at the first base of the codons for amino acids positions 5 and 7 of the heptad (see below). The formation of a self-assembling of structural material is mediated by the artificial recognition sequences (e.g. present sequence). A method of producing a self-assembling protein based structural material comprises mixing a multiplicity of 2 different artificial peptide sequences, designed so as to spontaneously heterodimerise. Both artificial peptides have a coiled-coil alpha helical secondary structure having at least 2 heptad units. The artificial peptides spontaneously heterodimerise into a self-assembled corrective associations as directural material. The method is used to produce nanoscale structural materials designed for specific functions e.g. membranes, fibres, absorbants, reactive meterials, etc. A specific corrective gloves and coated fabrics for fuel handlers and other hazardous materials of medical implants, wound healants and coated coated fabrics for fuel handlers wound cheat and coated coated fabrics for fuel handlers wound other companies.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sequence; self-assembling; protein-based structural material; heterodimerisation; coiled-coil alpha helix; permeable membrane; coated fabric.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW43024-28 represent recognition sequences obtained from the library of recognition sequence coded for by AAV04943. This DNA sequence is obtained by reverse translation. The codons were chosen to maximise
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Production of peptide hetero:dimer(s) - in the fabrication of self-assembling protein-based structural material(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56.5%; Score 213; DB 19; Length 84; llarity 93.6%; Pred. No. 2.7e-15; Conservative 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GDLKNKVAQLKRKVRSLKDKAAELKQHVSRLENEIEDLKAKIGDLNN 62
Disclosure; Fig 4C; 17pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Artificial recognition sequence 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Fig 4D; 17pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW43027 standard; Peptide; 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    95US-0452592.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                93US-0068948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              McGrath KP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1998-120029/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ψ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAY-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Recognition
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  selectively
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                spontaneous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US5712366-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27-JAN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kaplan DL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW43027;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW4302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
```

a

à

Similarity

84 AA;

```
ö
                                                                                                                                                                                                                                                                                               Coiled coil protein; copolymer blocker; vulnerary; wound treatment; integrin; fibroblast cell; matrix formation; skin regeneration; scaffold; antibiotic; dressing; abrasion; burn; Bl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This invention describes a novel synthetic block copolymer XYZ comprising two alpha-helical protein blocks X and Z, a water soluble, random-coil protein block Y which links X and Z, and linker proteins for linking the alpha-helical protein blocks to the random-coil protein block. The products of the invention have vulnerary activity. A copolymer produced from a gene that encodes a derivative of AC_l_0A copolymer was suspended
                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic block copolymer with an antibiotic compound, useful for dressing abrasion, burn or non-puncture wound, comprises two alpha-helical protein blocks, random-coil protein block and linker
                                   ö
54.4%; Score 205; DB 19; Length 84; 89.4%; Pred. No. 1.9e-14; tive 4; Mismatches 1; Indels
                                                                          16 GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Example 1; Column 25-26; 53pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Petka WA, McGrath KP, Tirrell DA;
                                                                                                                                                                       AAB10558 standard; protein; 42
                                                                                                                                                                                                                                                                                                                                                                                                                                                         97US-0956307.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-0956307.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (UYMA-) UNIV MASSACHUSETTS.
                                                                                                                                                                                                                                    (first entry)
                                Conservative
                                                                                                                                                                                                                                                                  Coiled coil protein B1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2000-542235/49.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       22-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                             Unidentified
                                                                                                                                                                                                                                    15-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                           US6090911-A.
                           42;
                                                                                                                                                                                                                                                                                                                                                                                                                         18-JUL-2000
                                                                                                                                                                                                      AAB10558;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               proteins
```

```
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
         δλ
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                in water to form a gel, which was then used to treat a wound. The cell binding domain was the integrin ArgGlyAspSer sequence which is known to bind to gpIIIA/IIIa proteins expressed on fibroblast cells necessary for matrix formation for the regeneration of skin. Fibroblasts become contrapped within the gel and thus remain at the site of the wound, serving as a scaffold for the regeneration of tissue surrounding the wound. A synthetic block copolymer together with an antibiotic compound is useful for wound dressing. The copolymer is useful for dressing abrashon, burn or non-puncture wound. Unlike conventional polymers, the polymers of equal size are produced biologically from a single template. The intermolecular binding of alpha-helical blocks are monodisperse which city is a uniform pore size depending on the length of random-coil block. The new copolymers have low molecular weight hence have decreased viscosity solutions or gels under suitable conditions. This sequence
                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Recognition sequence; self-assembling; protein-based structural material; spontaneous heterodimerisation; colled-coil alpha helix; selectively permeable membrane; coated fabric.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW43024-28 represent recognition sequences obtained from the library of recognition sequenes coded for by AAV04943. This DNA sequence is obtained by reverse translation. The codons were chosen to maximise expression in Escherichia coli, and to introduce useful restriction for subsequent genetic manipulations. The DNA sequence use a "mixed site" approach at the first base of the codons for amino acids positions 5 and 5 of the heptad (see below). The formation of a self-assembling structural material is mediated by the artificial recognition sequences (e.g. present sequence). A method of producing a self-assembling protein-based structural material comprises mixing a multiplicity of
                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Production of peptide hetero:dimer(s) - in the fabrication of self-assembling protein-based structural material(s)
                                                                                                                                                                                                                                                                                                                                                                           Length 42;
                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 SGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAK 42
                                                                                                                                                                                                                                                                                                                                                                       Score 199; DB 21;
Pred. No. 3.7e-14;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    SGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Artificial recognition sequence 5.
                                                                                                                                                                                                                                                                                                                                                         52.8%; Scu-
100.0%; Pre-
0; }
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Fig 4E; 17pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAW43028 standard; Peptide; 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95us-0452592.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             93US-0068948.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                               method of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Kaplan DL, McGrath KP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1998-120029/11.
                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                   42 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-MAY-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US5712366-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27-JAN-1998
                                                                                                                                                                                                                                                                                                                                                                                                               42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic.
                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAW43028;
                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAW43028
```

g

δ

8\$8888888888888888888888

```
AAW43024-28 represent recognition sequences obtained from the library of recognition sequenes coded for by AAV04943. This DNA sequence is obtained by reverse translation. The codons were chosen to maximise expression in Escherichia coil, and to introduce useful restriction for subsequent genetic manipulations. The DNA sequence use a "mixed site" subsequent genetic manipulations. The DNA sequence use a "mixed site" of the first base of the codons for amino acids positions 5 and 7 of the heptad (see below). The formation of a self-assembling structural material is mediated by the artificial recognition sequences c. 7 of the heptad sequence. A method of producing a self-assembling protein-based structural material comprises mixing a multiplicity of 2 different artificial peptide sequences, designed so as to spontaneously heterodimerise. Both artificial peptides about artificial peptides spontaneously heterodimerise into a self-assembled artificial peptides spontaneously heterodimerise into a self-assembled protein-based structural material. The method is used to produce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
2 different artificial peptide sequences, designed so as to spontaneously heterodimerise. Both artificial peptides have a coiled-coil alpha helical secondary structure having at least 2 heptad units. The artificial peptides spontaneously heterodimerise into a self-assembled protein-based structural material. The method is used to produce nanoscale structural materials designed for specific functions e.g. membranes, fibres, absorbants, reactive materials, etc. A specific application is in the fabrication of selectively permeable membranes for protective gjoves and coated fabrics for fuel handlers and other hazardous materials. Materials for medical implants, wound healants and other medical treatments would also be feasible using this technology.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Recognition sequence; self-assembling; protein-based structural material; spontaneous heterodimerisation; coiled-coil alpha helix; selectively permeable membrane; coated fabric.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Production of peptide hetero:dimer(s) - in the fabrication of self-assembling protein-based structural material(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16 GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2 GDLENEVAQLEREVRSLEDEAAELKQEVSRLENEIEDLKAKIGDLEN 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 197; DB 19;
Pred. No. 1.3e-13;
5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Artificial recognition sequence 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 52.3%; Scur
85.1%; Pred
6; /
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Disclosure; Fig 4A; 17pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW43024 standard; Peptide; 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95US-0452592.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             93US-0068948.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-MAY-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   McGrath KP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (USSA ) US SEC OF ARMY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1998-120029/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-MAY-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US5712366-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27-JAN-1998.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kaplan DL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAW43024;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
```

```
This invention describes a novel synthetic block copolymer XYZ comprising two alpha-helical protein blocks X and Z, a water soluble, random-coil protein block Y which links X and Z, and linker proteins for linking the alpha-helical protein blocks to the random-coil protein block. The calpha-helical protein blocks to the random-coil protein block. The products of the invention have vulnerary activity. A copolymer produced from a gene that encodes a derivative of AC_1_OA copolymer produced in water to form a gel, which was then used to treat a wound. The cell bluding domain was the integrin ArgGlyAspSer sequence which is known to compute of gpila/Tila proteins expressed on fibroblast cells necessary for matrix formation for the regeneration of sine. Fibroblasts become carrix formation for the regeneration of sisse surrounding the cwound. A synthetic block copolymer remain at the site of the wound, cound dressing. The copolymer is useful for dressing compound is useful for wound dressing. The copolymer is useful for dressing compound companies of equal size are produced biologically from a single template. The intermolecular binding of alpha-helical blocks are monodisperse which calves a uniform pore size depending on the length of random-coil block.
                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Coiled coil protein; copolymer blocker; vulnerary; wound treatment; integrin; fibroblast cell; matrix formation; skin regeneration; scaffold; antibiotic; dressing; abrasion; burn; Al.
                                          for
                                                                                healants and
nanoscale structural materials designed for specific functions e.g. membranes, fibres, absorbants, reactive materials, etc. A specific application is in the fabrication of selectively permeable membranes for protective gloves and coated fabrics for fuel handlers and other hazardous materials. Materials for medical implants, wound healarts and other medical treatments would also be feasible using this technology.
                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              dressing abrasion, burn or non-puncture wound, comprises two alpha-helical protein blocks, random-coil protein block and linker
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic block copolymer with an antibiotic compound, useful for
                                                                                                                                                                                                              ö
                                                                                                                                                                         84;
                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                   Length
                                                                                                                                                                                                                                              16 GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNN 62
                                                                                                                                                                       Score 181; DB 19;
Pred. No. 6.7e-12;
                                                                                                                                                                                                           10; Mismatches
                                                                                                                                                                     Score 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example 1; Column 25-26; 53pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Tirrell DA;
                                                                                                                                                                                                                                                                                                                                                                             AAB10557 standard; protein; 42 AA.
                                                                                                                                                                       48.08;
                                                                                                                                                                                           76.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97us-0956307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97us-0956307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (UYMA-) UNIV MASSACHUSETTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                     Local Similaricy
hes 36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Coiled coil protein Al.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    McGrath KP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2000-542235/49
                                                                                                                                  84 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US6090911-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Petka WA,
                                                                                                                                                                                                                                                                                                                                                                                                                  AAB10557;
                                                                                                                                  Sequence
                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Coiled
                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                         RESULT 10
AAB10557
                                                                                                                                                                                                           Matches
 8888888888
                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                 8
```

```
ö
                                                                                                                                                                                                                                                                            sequence; self-assembling; protein-based structural material;
heterodimerisation; coiled-coil alpha helix;
permeable membrane; coated fabric.
The new copolymers have low molecular weight hence have decreased viscosity solutions or gels under suitable conditions. This sequence represents the coiled coil protein Al which is used to illustrate the method of the invention.
                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                           Lys"
                                                                                                                                                                                                                                                                                                                                                                                                                                Lys"
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lys"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lys"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lys"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lys"
                                                                                                                                                                                                                                                                                                                                                              'note= "any chrged residue, preferably Glu or Lys"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        preferably Glu or Lys"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  preferably Glu or Lys"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      preferably Glu or Lys"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note= "any chrged residue, preferably Glu or Lys"
                                                                                                 ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            or
                                                                                                                                                                                                                                                        Amino acid sequence giving a library of recognition sequences.
                                                                                                                                                                                                                                                                                                                                                                                    'note= "any chrged residue, preferably Glu or
                                                                                                                                                                                                                                                                                                                                                                                                                             "any chrged residue, preferably Glu or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            preferably Glu or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note= "any chrged residue, preferably Glu or
                                                                                                                                                                                                                                                                                                                                                                                                                                                   preferably Glu or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Glu
                                                                           Length 42;
                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                          preferably
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          preferably
                                                                                                                               15 SGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAK 56
                                                                                               ò
                                                                           DB 21;
                                                                                     Pred. No. 6.4e-10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                 'note- "any chrged residue,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        residue,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            'note= "any chrged residue,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "any chrged residue,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'note= "any chrged residue,
                                                                                                                                                                                                                                                                                                                                                                                                           residue,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          residue,
                                                                          Score 159;
                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                         chrged
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "any chrged
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          chrged
                                                                                                                                                                                            Ź
                                                                                               10;
                                                                                                                                                                                          AAW43029 standard; Peptide; 42
                                                                        42.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "any
                                                                                                                                                                                                                                                                                                                                                                                                       /note= "any
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     93US-0068948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95US-0452592
                                                                                                                                                                                                                                    (first entry)
                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                              /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                McGrath KP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'note-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (USSA ) US SEC OF ARMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1998-120029/11.
                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 19
                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 17
                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 33
                                                                                                                                                                                                                                                                                                                                                     Misc-difference 5
                                                     42 AA;
                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAV04943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-MAY-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-MAY-1993;
                                                                                                                                                                                                                                   08-MAY-1998
                                                                                                                                                                                                                                                                              Recognition
                                                                                                                                                                                                                                                                                        spontaneous
                                                                                                                                                                                                                                                                                                 selectively
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US5712366-A.
                                                                                               32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27-JAN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Kaplan DL,
                                                                                                                                                                                                                                                                                                                       Synthetic.
                                                      Sequence
                                                                                                                                                                                                               AAW43029;
                                                                          Query Match
                                                                                      Local
                                                                                               Matches
                                                                                                                                                                               AAW43029
  8
                                                                                                                                                                                                               ó
                                                                                                                                      셤
```

```
Sequence
                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                      AAB08383
ID AABC
  ð
                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                        The present sequence is used to obtain an entire library of recognition sequences. The DNA sequence encoding the present sequence is obtained by reverse translation. The codons were chosen to maximize expression in Escherichia coli, and to introduce useful restriction for subsequent genetic manipulations. The DNA sequence use a "mixed site" approach at the first base of the codons for amino acids positions 5 and 7 of the heptad (see below). The formation of a self-assembling structural material is mediated by the artificial recognition sequences obtained from the present sequence. A method of producing a self-assembling protein-based structural material peptide sequences, designed so as to spontaneously content artificial peptide sequences, designed so as to spontaneously enterodimerise. Both artificial peptides have a colled-coil alpha helical secondary structure having at least 2 heptad units. The artificial peptides spontaneously heterodimerise into a self-assembled protein-based structural material. The method is used to produce nanoscale structural materials designed for specific functions e.g. membranes, fibres, absorbants, reactive materials, etc. A specific application is in the coated fabrics for fuel handlers and other mazardous materials. Materials for medical implants, wound healants and other medical treatments would selectively by.
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Binding partner; protein modification; post-translational modification; modulator; colled-coil structure.
                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                          ö
Production of peptide hetero:dimer(s) - in the fabrication of self-assembling protein-based structural material(s)
                                                                                                                                                                                                                                                                                                                                                  Length 42;
                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "fluorophore attachment site"
                                                                                                                                                                                                                                                                                                                                                 Score 124; DB 19;
Pred. No. 3.3e-06;
0; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                                                                                                                                                                                                                                                                                                                                                                              /note= "thrombin cleavage site"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Amino acid sequence of a coiled-coil peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mezna M;
                                    Claim 2; Columns 13-14; 17pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB08364 standard; peptide; 43 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Maschio A,
                                                                                                                                                                                                                                                                                                                                                 Query Match 32.9%;
Best Local Similarity 72.5%;
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25-FEB-2000; 2000WO-GB00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99GB-0004398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (FLUO-) FLUORESCIENCE LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Craig RK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2000-572119/53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 39
                                                                                                                                                                                                                                                                                                                           42 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               40200050902-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cleavage-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Jnidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-FEB-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31-AUG-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Colyer J,
                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB08364;
                                                                                                                                                                                                                                                                                                                                                                                                16
                                                                                                                                                                                                                                                                                                                                                                                                                       ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB08364
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The specification describes a method for analysing a sample. The method comprises immobilishing a polypeptide to a physical support, contacting the immobilised polypeptide with a test sample which may contain an agent capable of modifying the immobilised polypeptide, contacting the immobilised polypeptide with a binding partner polypeptide, where association of both polypeptide is dependent on the modification of state of the immobilised polypeptide, and measuring the association the binding partner polypeptide. The the binding partner polypeptide to the immobilised polypeptide.
                                                                                                                                                                                             The specification describes a method for analysing a sample. The method comprises immobilising a polypeptide to a physical support, contacting the immobilised polypeptide with a test sample which may contain an
                                                                                                                                                                                                                                                                                                               agent capable of modifying the immobilised polypeptide, contacting the immobilised polypeptide with a binding partner polypeptide, where association of both polypeptide is dependent on the modification state of the immobilised polypeptide, and measuring the association of the binding partner polypeptide, and measuring the association of the binding partner polypeptide to the immobilised polypeptide. The polypeptides is support and methods can be used to analyse a sample to determine if modification of a polypeptide is taking place and to identify modulators of the modification. This is useful for monitoring the post-translational modification of proteins. ABB08364 of sepresent binding partners which have collect-coll structures, and may be used in the method of the invention to assay for thrombin activity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Binding partner; protein modification; post-translational modification; modulator; colled-coil structure.
High throughput assay for monitoring modification of polypeptides and modulation of the modifications
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              for monitoring modification of polypeptides and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Peptide used to assay for tobacco etch virus protease activity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                32.1%; Score 121; DB 21; Length 43; 46.7%; Pred. No. 7e-06; tive 14; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 HHHHHHGGIA--QLEQEIAQLEQENRQLEQEIAQLEQEIAKLEQE 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 HHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENE 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mezna
                                                                                                                 Disclosure; Page 61; 128pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example 9; Page 80; 128pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB08383 standard; peptide; 43 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Maschio A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          High throughput assay for monito
modulation of the modifications
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-FEB-2000; 2000WO-GB00669.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99GB-0004398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 32.1%
Best Local Similarity 46.7%
Matches 21; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (FLUO-) FLUORESCIENCE LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Craig RK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4PI; 2000-572119/53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    43 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200050902-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-FEB-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31-AUG-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Colyer J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB08383;
```

VASREEVWSLESEVSSLESRVSSLRDEVSSL 92

62

93

AAB70796 standard; Protein;

æ

8888888

셤

à

```
g
                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                          ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This invention describes a novel protease substrate for detecting a target protease which comprises a polypeptide containing a high-affinity binding site (1) for binding to a support, a cleavage site (2) specific for the target protease, and a sequence of amino acids forming a stretched helix (3). The substrate is useful for measuring protease activities and in online screening assays for proteases or protease inhibitors, especially based on surface plasmon resonance.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                online screening assays for proteases or protease inhibitors, comprises a polypeptide with a binding site, a cleavage site and a stretched
polypeptides, support and methods can be used to analyse a sample to determine if modification of a polypeptide is taking place and to identify modulators of the modification. This is useful for monitoring the post-translational modification of proteins. AAB08383-84 represent binding partners which are used, in the method of the invention, to assay for tobacco etch virus (TEV) protease activity.
                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---LKNKVAQLKRKVRSLKDKAAELKQEVSRLEN- 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protease substrate useful for measuring protease activities and in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23;
                                                                                                                                                                                        ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Protease substrate; detection; cleavage site; stretched helix;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 103;
                                                                                                                                                      DB 21; Length 43;
                                                                                                                                                                                      8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20; Indels
                                                                                                                                                                                                                   HHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENE 49
                                                                                                                                                                                                                                    |||||||| :| |: ::|||:::|| |: HHHHHHGGIA--QLEQEIAQLEQEIAQLEQEIAXLEQE 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31.7%; Score 119.5; DB 22;
31.9%; Pred. No. 2.7e-05;
ive 19; Mismatches 20; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Steinruecke P;
                                                                                                                                                    ; Score 121; DB 21
; Pred. No. 7e-06;
14; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (IMBM-) IMB INST MOLEKULARE BIOTECHNOLOGIE EV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----EIEDLKAKIGDL-NNTSGIRRPAAKL 73
                                                                                                                                                                                                                                                                                                                               AAB70797 standard; Protein; 103 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                  HIV protease-associated protein #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hill O, Hillsch A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Fig 2; 14pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human immunodeficiency virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99DE-1041239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99DE-1041239
                                                                                                                                        Query Match
Best Local Similarity 46.7%
                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 31.9
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HHHHHHGSMASGD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-246019/26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   103 AA;
                                                                                                                     43 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DE19941239-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Aldinger U,
                                                                                                                                                                                                                                                                                                                                                                                                 25-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   26-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     screening
                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                   AAB70797;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                49
                                                                                                                                                                                                                                                                                                                   AAB7079
```

ò ద ö

```
This invention describes a novel protease substrate for detecting a target protease which comprises a polypeptide containing a high-affinity binding site (1) for binding to a support, a cleavage site (2) specific for the target protease, and a sequence of amino acids forming a stretched helix (3). The substrate is useful for measuring protease activities and in online screening assays for proteases or protease inhibitors, especially based on surface plasmon resonance.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Protease substrate useful for measuring protease activities and in online screening assays for proteases or protease inhibitors, comprises a polypeptide with a binding site, a cleavage site and a stretched
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 HHHHHHGSMAS----GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLEN-----EIEDLK 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11;
                                                                                                                                                               Protease substrate; detection; cleavage site; stretched helix;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 22; Length 93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        31.4%; Score 118.5; DB 22; 32.5%; Pred. No. 3.1e-05; iive 20; Mismatches 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hillsch A, Steinruecke P;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (IMBM-) IMB INST MOLEKULARE BIOTECHNOLOGIE EV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 December 18, 2002, 16:17:27
                                                                                                         protease-associated protein #1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Fig 2; 14pp; German.
                                                                                                                                                                                                                                            Human immunodeficiency virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           55 AKIGDL-NNTSGIRRPAAKL 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEVSSLESRVSSLRDEVSSL 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99DE-1041239.
                                                                                                                                                                                                                                                                                                                                                                                                              99DE-1041239
                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              H111 O,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2001-246019/26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search completed: Decemb
Job time : 5.94091 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAF61357
                                                                                                                                                                                                                                                                                                  DE19941239-A1
                                                                                                                                                                                                                                                                                                                                                                                                              26-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26-AUG-1999;
                                                   25-MAY-2001
                                                                                                                                                                                                                                                                                                                                                        01-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Aldinger U,
AAB70796;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              63
```

3;

_

```
7, Appli
9, Appli
12, Appli
10, Appli
11, Appli
11, Appli
12, Appli
20, Appli
20, Appli
21, Appli
22, Appli
22, Appli
31, Appl
                                                                                                                                                                        December 18, 2002, 16:14:14; Search time 1.70376 Seconds (without alignments) 1277.936 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 18, Appl
Sequence 17, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 17, Sequence 8, Sequence 9, Sequence 20, Sequence 17, Sequence 11, Sequence 11, Sequence 12, Sequence 22, Sequence 22, Sequence 23, Sequence 14, Sequence 15, Sequence 11, Sequence 11, Sequence 11, Sequence 21, Sequenc
                                                                                                                                                                                                                                                                                                                                                               74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                                                                         1 MRGSHHHHHHGSMASGDLKN.....AKIGDLNNTSGIRRPAAKLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1: /cgn2_6/ptodate/1/laa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/laa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/laa/6A_COMB.pep:*
5: /cgn2_6/ptodata/1/laa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/laa/pcTUS_COMB.pep:*
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-956-307B-18

US-08-956-307B-17

US-08-452-592B-9

US-08-452-592B-9

US-08-452-592B-9

US-08-452-592B-10

US-08-452-592B-10

US-08-452-592B-10

US-08-955-307B-1

US-09-329-884-18

US-09-329-884-12

US-09-329-884-12

US-09-329-884-12

US-09-329-884-12

US-09-329-884-14

US-09-329-884-14

US-09-329-884-14

US-09-329-884-14

US-09-329-884-14

US-08-956-307B-4

US-08-956-307B-4

US-08-956-307B-4

US-08-956-307B-1

US-08-956-307B-1

US-08-956-307B-1

US-08-956-307B-1

US-08-956-307B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCT-US94-06655-9
US-08-457-245-21
US-09-299-495F-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            262574 seqs, 29422922 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                       BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                 US-09-490-291-11
377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.5
91.5
91
90
87
                                                                                                                                                                                                                                                                                                                                                                                                                       Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Minimum DB 8
Maximum DB 9
                                                                                                                   OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database
                                                                                                                                                                                 Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š.
```

Sequence 12, Appl. Sequence 10, Appl. Sequence 10, Appl. Sequence 5383, Appl. Sequence 37, Appl. Sequence 37, Appl. Sequence 39, Appl. Sequence 39, Appl. Sequence 21, Appl. Sequence 24, Appl. Sequence 26, Appl. Sequence 27, Appl. Sequence 27, Appl. Sequence 28, Appl. Sequence 29, Appl. Sequence 29, Appl.		Length 76; Indels 5; Gaps 1; QEVSRLENEIEDLKA 55
28 81 21.5 126 1 US-08-268-348A-12 30 81 21.5 133 1 US-08-268-348A-8 30 81 21.5 133 1 US-08-268-348A-10 31 80.5 21.4 202 4 US-09-134-001C-5383 32 80.5 21.4 1587 4 US-09-100-094-46 33 79.5 21.1 31 2 US-08-62-227-37 34 79.5 21.0 32 3 US-08-77-336-1 36 79 21.0 39 2 US-08-77-36-39 37 79 21.0 39 2 US-08-67-865-39 38 79 21.0 39 4 US-09-263-975-39 40 79 21.0 69 2 US-08-687-865A-21 41 79 21.0 178 4 US-09-011A-11 42 79 21.0 178 4 US-09-101A-11 43 79 21.0 254 4 US-09-004-731-89 45 79 21.0 254 4 US-09-004-731-89	RESULT 1  US-08-956-3078-18  Sequence 18, Application US/089563078  Patent No. 6090311  SEREAL INFORMATION: APPLICANT: Petka, Wendy A. APPLICANT: ROWIND REVERSIBLE HYDROGELS NUMBER OF SEQUENCES: 24 CORRESPONDENCES: 24 CORRESPONDENCES: 18 Richardson P.C. STREET: 225 Franklin Street CITY: BOSTON STREET: 225 Franklin Street CITY: BOSTON STREET: 125 Franklin Street CONPUTRY: USAA COMPUTR: USAA COMPUTER: IBM COMPATION MEDIUM TYPE: Diskette COMPUTER: IBM COMPATION APPLICATION DATA: APPLICATION NUMBER: 12-0CT-1997 PRIOR APPLICATION DATA: APPLICATION NUMBER: 32-0CT-1997 PRIOR APPLICATION NUMBER: 32-0CT-1997 TELEPONE: FRASE, J. PRECE SEGUENCE OF THE SEGUENCE OF T	Query Match 74.1%; Score 279.5; DB 3; Length 76; Best Local Similarity 89.1%; Pred. No. 2.3e-23; Matches 57; Conservative 1; Mismatches 1; Indels 5; G Qy 1 MRGSHHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA

```
COUNTRY: United States of America ZIP: 01760-5035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Sequence 8, Application US/08452592B
; Patent No. 5712366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: 508-233-5167
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
    NUMBER OF SEQUENCES: 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; TYPE: amino acid
; TOPOLOGY: linear
US-08-452-5928-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 4
US-08-452-592B-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ä
1 MRGSHHHHHHGSDDDDKWASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLKNEIEDLKA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MRGSHHHHHHGS-----MASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 7, Application US/08452592B;
Patent No. 5712366
GENERAL INFORMATION:
APPLICANT: McGrath, Kevin P.
TITLE OF INVENTION: Fabrication of Nanoscale Materials Using TITLE OF INVENTION: Self-Assembling Proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 64.6%; Score 243.5; DB 3; Best Local Similarity 75.0%; Pred. No. 1.7e-19; Matches 48; Conservative 10; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IDM COMPALIALLE
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSEQ for Windows version 2.0b
CURREWT APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/08/956,307B
FILING DATE: 22-0CT-1997
FILING DATE:
FILING DATE:
                                                                                                                                                                        Sequence 17, Application US/08956307B
Patent No. 6090911
GENERAL INFORMATION:
APPLICANT: Tirrell, David A.
APPLICANT: Tirrell, David A.
APPLICANT: Tirrell, David A.
APPLICANT: Revin P. McGrath
TITLE OF INVENTION: REVERSIBLE HYDROGELS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               07880/033001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY AGENT INFORMATION:
NAME: FESSE, J. PETER
REGISTRATION NUMBER: 32,983
REFRENCE/DOCKET NUMBER: 0788
TELEPHONE: 617/542-5070
TELEPHONE: 617/542-9906
TELERAX: 200154
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 aning acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: USA
ZIP: 02110-2804
                                                                                                                                       RESULT 2
US-08-956-307B-17
                                      56 KIGD 59
                                                                   ||||
61 KIGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56 KIGD 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : | | |
61 EIGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-956-307B-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 3
US-08-452-592B-7
                                                                                                                                                                                                                                                                                                                                                                                                          CITY:
STATE:
g
                                                                                윱
                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:

APPLICANT: McGrath, Kevin P.

APPLICANT: McGrath, Revin P.

APPLICANT: McGrath, David L.

TITLE OF INVENTION: Self-Assembling Proteins

TITLE OF INVENTION: Self-Assembling Proteins

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: U.S. Army Soldier Systems Command

STREET: AMSC-CC (Patent Counsel)

CITY: Massachsuetts
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 GDLKNKVAQLKRKVRSLKDRAAELKQEVSRLENEIEDLKAKIGDLKN 48
COMPUTER TAMSCC-CC (Patent Counsel)
STREET: AMSCC-CC (Patent Counsel)
CITY: NALICK
CITY: MASCC-CC (Patent Counsel)
CITY: MASCC-CC (Patent Counsel)
CITY: MASSCC-CC (Patent Counsel)
COUNTRY: United States of America
ZIP: 01760-5035
COMPUTER: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: IBM PC Compatable
COMPUTER: IBM PC COMPATA:
APPLICATION TAMER: US/08/452,592B
FILING DATE: 25 May 1995
CLASSIFICATION DATA:
APPLICATION NUMBER: US/08/452,592B
FILING DATE: 25 May 1993
ATTORNEY/AGET INFORMATION:
NAME: Lamming, John H.
NAME: Lamming, John H.
NAME: Lamming, Vincent J.
REGISTRATION NUMBER: 34,857 (Lamming)
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEFRAX: S08-233-4510
TELEFRAX: S08-233-4510
TELEFRAX: S08-233-4510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16 GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage COMPUTER: IBM PC Compatable OPERATING SYSTEM: DOS 6.22 SOFTWARE: WordPerfect Version 6.1 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/452,592B FILING DATE: 25 May 1995 CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
58.6%; Score 221; DB 1;
Best Local Similarity 97.9%; Pred. No. 4.9e-17;
Matches 46; Conservative 0; Mismatches 1
```

ö

ö

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                      Length 84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Indels
                                                                                                                                                                                                                                                                                         16 GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 SGDLKNKVAQLKRKVRSLKDKAABLKQEVSRLENEIEDLKAK 42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: DISKette
COMPUTER: IBM Compatible
COMPUTER: IBM Compatible
COMFURING SYSTEM: Windows 95
SOFFWARE: FastSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,307B
FILING DATE: 22-0CT-1997
PRIOR APPLICATION DATA:
FILING DATE: P2-0CT-1997
APPLICATION NUMBER:
FILING DATE: PAPER: FILING DATE:
FILING DATE: FASSE, J. Peter
NAME: Fasse, J. Peter
NAME: Fasse, J. Peter
NAME: Fasse, J. Peter
NAME: Fasse, J. Peter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15 SGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENETEDLKAK
                                                                                                                                Score 205; DB 1;
Pred. No. 2.6e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 199; DB 3;
Pred. No. 5e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 52.8%; Score 100.0%; Pred. No. con 100.0%; Pred. No. con 100.0%; Mismatches
                                                                                                                                                                                           4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: Pecks, Wendy A.
APPLICANT: Tirrell, David A.
APPLICANT: Kevin P. McGrath
TITLE OF INVENTION: REVERSIBLE HYDROGELS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         07880/033001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: Fish & Richardson P.C. STREET: 225 Franklin Street CITY: Boston STATE: MA COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 10, Application US/08452592B
Patent No. 5712366
GENERAL INFORMATION:
APPLICANT: McGrath, Kevin P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 2, Application US/08956307B Patent No. 6090911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE DOCKET NUMBER: 0781
TELEPHONE: 617/542-5070
TELEPHONE: 617/542-8906
TELERA: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 aming acids
                                                                                                                                54.48;
                                                                                                                                                      Best Local Similarity 89.49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ) MOLECULE TYPE: peptide US-08-956-307B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
            amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      inear
                                           TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                02110-2804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-452-592B-10
                                                                 US-08-452-592B-9
                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-956-307B-2
                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 7
                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 9, Application US/08452592B
Patent No. 5712366
Paper No. 5712366
Patent No. 571236
Patent No. 5712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56.5%; Score 213; DB 1; Length 84; 93.6%; Pred. No. 3.6e-16; Live 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET: Autick
CITY: Natick
STATE: Massachuetts
COUNTRY: United States of America
ZIP: 01760-5035
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: IBM PC Compatable
OPERATING SYSTEM: DOS 6.22
SOFTWARE: WordPerfect Version 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,592B
FILING DATE: 25 May 1995
FILING DATE: 25 May 1995
FILING DATE: 25 May 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16 GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 GDLKNKVAQLEREVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLKN 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICALLON

FILING DATE: 25 May 1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/068,948

FILING DATE: 25 May 1993

ATTORNEY/AGENT INFORMATION:
NAME: Lamming, John H.
NAME: Ranucci, Vincent J.
REGISTRATION NUMBER: 22,062 (Donahue)

REGISTRATION NUMBER: 34,857 (Lamming)

REGISTRATION NUMBER: 32,9579 (RANUCCI)

REGISTRATION NUMBER: 32,9579 (RANUCCI)
                                                                                                                NAME: Donahue, Richard J.

NAME: Lamming, John H.

NAME: Ranucci, Vincent J.

REGISTRATION NUMBER: 22,062 (Donahue)
REGISTRATION NUMBER: 34,857 (Lamming)
REGISTRATION NUMBER: 29,579 (Ranucci)
REFERENCE/DOCKET NUMBER: NA-1096D
TELECOMUNICATION INFORMATION:
TELEPHONE: 508-233-4510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: NA-1096D TELECOMMUNICATION INFORMATION: TELEPHONE: 508-233-4510
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/068,948
FILING DATE: 25 May 1993
ATTORNEY/AGENT INFORMATION:
NAME: Donahue, Richard J.
                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: 508-233-5167 INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: 508-233-5167
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                508-233-5167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid;
TOPOLOGY: linear
US-08-452-5928-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-452-592B-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

ö

```
Length 84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 GDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAEIGDLEN 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16 GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
COMPUTER: IBM Compatible
COMPUTER: FastSEM: Windows 95
SOFTWARE: FastSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,307B
FILNG APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 48.0%; Score 181; DB 1; 76.6%; Pred. No. 9.7e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
APPLICANT: Petks, wendy A.
APPLICANT: Tirrell, David A.
APPLICANT: Kevin P. McGrath
TITLE OF INVENTION: REVERSIBLE HYDROGELS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
            ETLING DATE: 25 May 1995
CLASSIETCATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/068,948
FILING DATE: 25 May 1993
ATTORNEY/AGENT INFORMATION:
NAME: Donablue, Richard J.
NAME: Lamming, John H.
NAME: Ranucci, Vincent J.
REGISTRATION NUMBER: 29,579 (Ran REGISTRATION NUMBER: 29,579 (Ran RECISTRATION NUMBER: 39,579 (Ran RELECHONMUNICATION INFORMATION:
TELEPRAX: 508-233-4510
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E: Fish & Richardson P.C.
225 Franklin Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/08956307B
Patent No. 6090911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FASSE, J. PETER
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 0786
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 76.6%
Matches 36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: 225 Frank
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   amino acid
GY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; TOPOLOGY:
US-08-452-592B-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-956-307B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
TITLE OF INVENTION: Fabrication of Nanoscale Materials Using TITLE OF INVENTION: Self-Assembling Proteins NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 6, Application US/08452592B
Patent No. 5712366
GENERAL INFORMATION:
APPLICANT: McGrath, Kevin P.
APPLICANT: Kaplan, David L.
TITLE OF INVENTION: Fabrication of Nanoscale Materials Using TITLE OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   52.3%; Score 197; DB 1; Length 84;
85.1%; Pred. No. 1.9e-14;
tive 6; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16 GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 GDLENEVAQLEREVRSLEDEAAELKQEVSRLENEIEDLKAKIGDLEN 48
                                                                                                                                                                               STATE: MASSACKSUELTS

COUNTRY: United States of America
ZIP: 01760-2035

COUNTRY: United States of America
ZIP: 01760-2035

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: IBM PC Compatable
OPERATING SYSTE: DOS 6.2
SOFTWARE: DOS 6.2
SOFTWARE: MORDER: US/08/452,592B
FILING DATE: 25 May 1995
FILING DATE: 25 May 1995
CLASSIFICATION NUMBER: 08/068,948
FILING DATE: 25 May 1993
ATTORNEY/AGENT: NFORMATION:
NAME: Donahue, Richard J.
NAME: Lamming, John H.
NAME: Lamming, John H.
NAME: Ranucci, Vincent J.
REGISTRATION NUMBER: 32,657 (Lamming)
REGISTRATION NUMBER: 29,579 (Ranucci)
REFERENCE/DOCKET NUMBER: Na-1096D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSE: U.S. Army Soldier Systems Command STREET: AMSCC-CC (Patent Counsel)
CITY: Natiok
STATE: Massachsuetts
COUNTRY: United States of America
ZIP: 01760-5035
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKette, 3.50 inch, 2.0 MB storage COMPUTER: IBM PC COMPATEN SYSTEM: DOS 6.22
SOFTWARE: WordPerfect Version 6.1
CURRENT APPLICATION DATA:
                                                             STREET: Massart
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US/08/452,592B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-233-4510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 85.1
Matches 40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: 508-233-5167
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          amino acid
GY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-452-592B-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-452-592B-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

ô

0; Gaps

```
APPLICANT: Heinrikson, APPLICANT: Tomasselli,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
US-09-329-884-20
                                                                                                                                                                   SEQ ID NO 1
LENGTH: 286
                                                                                                                                                                                                          TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                             ö
                                                                                                                                                                                                                                                                                                               APPLICANT: McGrath, Kevin P.
APPLICANT: McGrath, Kevin P.
TITLE OF INVENTION: Fabrication of Nanoscale Materials Using NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSE: U.S. Army Soldier Systems Command
STREET: AMSCC-CC (Patent Counsel)
                                                        Length 42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 42;
                                                                                           0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: Massachuetts
COUNTRY: United States of America
ZIP: 01760-5035
COMPUTER READBLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: IBM PC Compatable
OPERATING SYSTEM: DOS 6.22
SOFTWARE: WordPerfect Version 6.1
CURRENT APPLICATION DATA:
APPLICATION MUBER: US/08/452,592B
FILING DATE: 25 May 1995
CLASSIFICATION: 530
                                                                                                                              15 SGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAK 56
                                                                                                                                                    1 SGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAE 42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16 GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                      42.2%; Score 159; DB 3; 76.2%; Pred. No. 9.9e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12.9%; Score 124; DB 1; 21.5%; Pred. No. 5.7e-07; 29; Conservative 0; Mismatches 11.
                                                                                           10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Donahue, Richard J.

NAME: Lamming, John H.

NAME: Ranucci, Vincent J.

REGISTRATION NUMBER: 22.062 (Donahue)

REGISTRATION NUMBER: 34,857 (Lamming)

REGISTRATION NUMBER: 29,579 (Ranucci)

REFERENCE, POCKET NUMBER: NA-1096D

TELECOMUNICATION INFORMATION:

TELEPHONE: 508-233-4510
                                                                                                                                                                                                                                                           Sequence 11, Application US/08452592B
Patent No. 5712366
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/068,948

FILING DATE: 25 May 1993

ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1, Application US/09360017 Patent No. 6303329 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: 11:
                                                                        Best Local Similarity 76.23
Matches 32; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               508-233-5167
 peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 29; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         amino acid
; MOLECULE TYPE:
US-08-956-307B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-452-592B-11
                                                                                                                                                                                                                                           US-08-452-592B-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
US-09-360-017-1
                                                        Query Match
                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
ä
                                                                                                                                                                                                                                                                                                                                                                                           ë
                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MRGSHHHHHHGSMASGD-----LKNKVAQLKRKVRSL----KDKAAELKQEVSR 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 HHHHHGGGGEVSALEKEVSALEKEVSALEKEVSALEKEVSALEKGGGGEFARAQLSERMT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16;
                                                                                                                                                                                                                                                                                                                                                                                        23;
                                                                                                                                                                                                                                                                                                                                          Length 286;
                                                of Procaspase
                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 175;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 25.5%; Score 96; DB 4; Length 175
Best Local Similarity 28.6%; Pred. No. 0.0031;
Matches 22; Conservative 15; Mismatches 24; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 HHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENE---
                                                                                                                                                                                                                                                                                                                                     Score 101.5; DB 4;
Pred. No. 0.0014;
8; Mismatches 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-129-884-18

Sequence 18, Application US/09329884

Patent No. 6342233

GENERAL INFORMATION:
APPLICANT: LIVIN, Randall T.
APPLICANT: Hodges, Robert S.
TITLE OF INVENTION: PSEUDOMONAS TREATMENT:
TITLE OF INVENTION: COMPOSITION AND METHOD:
FILE REFERENCE: 8900-0008.30
CURRENT APPLICATION NUMBER: US/09/329,884

CURRENT PILING DATE: 1999-06-11
EARLIER RPPLICATION NUMBER: US 60/089,155
EARLIER PILING DATE: 1998-06-12

NUMBER OF SEQ ID NOS: 22

SOFTWARE: FSELEG for Windows Version 3.0

SEQ ID NO 18

LENGTH: 175

TENGTH: 175
               APPLICANT: Koeplinger, Kenneth A.
TITLE OF INVENTION: Method for Autoactivation
TITLE REFERENCE: Docket No. 6303239 6172.N CP
CURRENT APPLICATION NUMBER: US/09/360,017
CURRENT FILING DATE: 1999-07-23
SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 20, Application US/09329884
Patent No. 634233
GENERAL INFORMATION:
APPLICANT: Irvin, Randall T.
APPLICANT: Hodges, Robert S.
TITLE OF INVENTION: PSEUDONONAS TREATMENT
TITLE OF INVENTION: COMPOSITION AND METHOD
FILE REFERENCE: 8900-0008.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; ORGANISM: Pseudomonas aeruginosa
US-09-329-884-18
Alfredo G
                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 34.6%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       46 LENEIEDLKAKIGDLNNT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 LHS-----IRDRNGT 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       50 -IEDLKAKIGDLNNTSG 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LASGLKTKVSDIFSQDG 80
                                                                                                                                                                                                                                                              ; ORGANISM: Homo sapiens
US-09-360-017-1
```

```
ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1;
                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
9
                                                                                                                                                                                                                                                                                                                                                                                  6 HHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                                                                                                                                                                                                                                                                                                                                                                                     4 HHHHHGG-----GGEIEALKAEIEALKAEIEALKAEIEALKA 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6 HHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 HHHHHGG-----GGEIEALKAEIEALKAEIEALKAEIEALKA 47
                                                                                                                                                                                                                                                                                          Query Match 24.9%; Score 94; DB 4; Length 168; Best Local Similarity 40.0%; Pred. No. 0.0049; Matches 20; Conservative 8; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 94; DB 4; Length 169
Pred. No. 0.0049;
8; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 12, Application US/09329884

Sequence 12, Application US/09329884

Patent No. 6342233

GENERAL INFORMATION:
APPLICANT: Irvin, Randall T.
APPLICANT: Irvin, Redges, Robert S.
TITLE OF INVENTION: COMPOSITION AND METHOD FILE REFERENCE: 8900-0008.30

CURRENT APPLICATION: NUMBER: US/09/329,884

CURRENT FILING DATE: 1999-66-11

EARLIER FILING DATE: 1998-66-12

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PASLESO for Windows Version 3.0

SEQ ID NO 12

LENGTH: 169

TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Irvin, Randall T.
APPLICANT: Hodges, Robert S.
TITLE OF INVENTION: DSEUDOWONAS TREATMENT
TITLE OF INVENTION: COMPOSITION AND METHOD
FILE REFERENCE: 8900-0008.30
CURRENT APPLICATION NUMBER: US/09/329,884
CURRENT FILING DATE: 1999-06-11
EARLIER APPLICATION NUMBER: US 60/089,155
EARLIER FILING DATE: 1998-06-12
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FESTSEQ for Windows Version 3.0
SEQ ID NO 16
LENGTH: 175
CURRENT APPLICATION NUMBER: US/09/329,884
CURRENT FILING DATE: 1999-06-11
EARLIER APPLICATION NUMBER: US 60/089,155
EARLIER FILING DATE: 1998-06-12
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 20
LENGTH: 168
TYPE: PRT
CREALISM: Pseudomonas aeruginosa
US-09-329-884-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-329-884-16
Sequence 16, Application US/09329884
Patent No. 6342233
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Pseudomonas aeruginosa
US-09-329-884-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 24.9%;
Best Local Similarity 40.0%;
Matches 20; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-329-884-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                          ð
```

```
ä
                                      Gaps
                                      9
Query Match 24.9%; Score 94; DB 4; Length 175; Best Local Similarity 40.0%; Pred. No. 0.0051; Matches 20; Conservative 8; Mismatches 16; Indels
                                                                                             6 HHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                                         ð
                                                                                                    g
```

Search completed: December 18, 2002, 16:23:15 Job time: 2.70376 secs

```
December 18, 2002, 16:21:10 ; Search time 1.02226 Seconds (without alignments) 1213.493 Million cell updates/sec
                                                                                                                                                           1 MRGSHHHHHHGSMASGDLKN.....AKIGDLNNTSGIRRPAAKLN 74
                                                                                                                                                                                                                                                                                                                                                                       GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                           106657 seqs, 16763532 residues
                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                    - protein search, using sw model
                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                 length: 0
length: 2000000000
                                                                                                                                 US-09-490-291-11
377
                                                                                                                                 Title:
Perfect score:
                                                                                                                                                                                                                                                                                Minimum DB seq
Maximum DB seq
                                                                                                                                                                                    Scoring table:
                                                   OM protein
                                                                                                                                                           Sequence:
                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                               Run on:
```

14: /cgn2_b/ptodata/l/pubpaa/US60_PUBCOMB.pep:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Sequence 2, Appli Sequence 1, Appli Sequence 1, Appli Sequence 20, Appli Sequence 12, Appli Sequence 12, Appli Sequence 14, Appli Sequence 46, Appli Sequence 46, Appli Sequence 14, Appli Sequence 51, Appli Sequence 62, Appli Sequence 64, Appli Sequence 67, Appli Description US-09-865-159-22 US-10-007-805-159-14 US-10-007-805-551 US-09-987-107-50 US-09-813-820-7 US-09-862-915-1 US-09-865-159-18 US-09-865-159-18 US-09-865-159-20 US-09-865-159-12 US-09-865-159-12 US-09-885-159-16 US-09-813-820-2 US-09-987-107-44 US-09-987-107-46 US-09-987-107-56 US-09-987-107-62 US-09-987-107-64 SUMMARIES Query Match Length DB 101.5 96 95.5 94 93.5 92 92 92 91.5 91.5 91.5 90 90 88 88 Score 

Sequence 60, Appl Sequence 44, Appl	Sequence 48, Appl Sequence 42, Appl	Sequence 2, Appl1	Sequence 66, Appl		Sequence 90, Appl	Sequence 37, App.	à	Sequence 4, Appl1	Sequence 8, Appl1	Sequence 2, Appli		Sequence 28, Appl	Sequence 9, Appli	Sequence 119, App	Sequence 119, App	Sequence 1, Appli	Sequence 3, Appl1				Sequence 217, App	Sequence 6, Appli	Sequence 3, Appl1
	9 US-09-987-107-48 10 US-09-732-091-42		_	9 US-09-987-107-54	9 US-09-870-759-90		10 US-09-884-681-39				10 US-09-813-820-6				10 US-09-874-923-119	_		9 US-09-987-107-52	9 US-09-987-107-68	10 US-09-845-157-2	۰	9 US-09-051-013-6	9 US-09-051-013-3
323 265	316 504	85	324	316	104	31	39	211	139	439	512	559	14	193	193	326	30	261	344	716	879	17	50
23.2	22.1 22.0	21.8	21.8	21.5	21.2	21.1	21.0	21.0	20.7	20.7	20.7	20.4	20.3	20.3	20.3	20.5	19.9	19.8	19.8	19.8	19.8	19.6	19.6
87.5 87	83.5 83	82	82	81	80	79.5	79	79	78	78	78	77	76.5	76.5	76.5	16	75	74.5	74.5	74.5	74.5	74	74
20	22 23	24	25	56	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

#### ALIGNMENTS

```
Sequence 2, Application US/09982992A

Sequence 2, Application US/09982992A

Patent No. US20020164337al

GENERAL INFORMATION:

TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE MAP PROTEIN AND METHOD OF USE IN

TITLE OF INVENTION: AND PREVENTING INFECTIONS

FILE REFERENCE: P06922US02/ABS

CURRENT APPLICATION NUMBER: 08/09/982,992A

CURRENT APPLICATION NUMBER: 60/277,287

PRIOR FILING DATE: 2001-03-21

PRIOR FILING DATE: 2001-03-21

PRIOR FILING DATE: 2001-03-20

NUMBER OF SEQ ID NOS: 6

SOFTWARE: Patentin version 3.0

SEQ ID NO 2

LENGTH: 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     27.6%; Score 104; DB 9; Length 131; llarity 30.7%; Pred. No. 0.0029; Conservative 9; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | |:
6.1 IDLRLSKQAKYTVHFKNGTKRVVDLKAGIHTADLINTSDIK 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               35 -----KAAELKQEVSRLENEIEDLKAKI--GDLNNTSGIR 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 1, Application US/09862915
Patent No. US2002000105051
GENERAL INFORMATION:
APPLICANT: Heinrikson, Robert I.
                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Staphylococcus aureus US-09-982-992A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 31; Conserv
US-09-982-992A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-862-915-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
```

```
PRIOR FILING DATE: 1999-12-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT
                                                                                                                                                                                   TYPE: PRT
                                                                                                                                                                                                                        FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 9, Application US/09742373
Patent No. US20020052471A1
GENERAL INFORMATION:
APPLICANT: Althaus, Harald
APPLICANT: Hauser, Hans-Peter
TITLE OF INVENTION: Human Procalcitonin and the Preparation and Use Thereof
FILE REFERENCE: 0552.1445.00
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 19962434.8
                                                                                                                                                                                                                                                                                                  ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ä
                                                                                                                                                                                                                                                                                                                                                          1 MRGSHHHHHHGSMASGD-----LKNKVAQLKRKVRSL-----KDKAAELKQEVSR 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 18, Application US/09865159
Patent No. US20020098196A1
GENERAL INFORMATION:
APPLICANT: ITVIO, Randall T.
APPLICANT: Hodges, Robert S.
TITLE OF INVENTION: PSEUDOMONAS TREATMENT
TITLE OF INVENTION: COMPOSITION AND METHOD
FILE REFERENCE: 8900-0008 30
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/329,884
PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-11
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 18
LENGTH: 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25.5%; Score 96; DB 10; Length 175; 28.6%; Pred. No. 0.021;
APPLICANT: Tomasselli, Alfredo G.
TITLE OF INVENTION: Method for Autoactivation of Procaspase FILE REFERENCE: Docket No. US20020045205A1 6172
CURRENT APPLICATION NUMBER: US/09/862,915
CURRENT FILING DATE: 2001-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO. 1286
                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 HHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENE-----
                                                                                                                                                                                                                                                          DB 10;
                                                                                                                                                                                                                                                        Score Lul., Pred. No. 0.011;
                                                                                                                                                                                                                                                                                              8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15; Mismatches
                                                                                                                                                                                                                                                          Score 101.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Pseudomonas aeruginosa
                                                                                                                                                                                                                                                        26.9%;
34.6%;
                                                                                                                                                                                                                                                                                                                                                                                                             46 LENEIEDLKAKIGDLNNT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                               61 LHS-----IRDRNGT 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 25.5%
Best Local Similarity 28.6%
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 LASGLKTKVSDIFSQDG 80
                                                                                                                                                                                                                                                                      Local Similarity 34.6
nes 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    50 -IEDLKAKIGDLNNTSG 65
                                                                                                                                                                           TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 3
US-09-865-159-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-865-159-18
                                                                                                                                                                                                                  US-09-862-915-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-742-373-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                            Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Irvin, Randall T.
APPLICANT: Irvin, Randall T.
APPLICANT: Irvin, Randall T.
TITLE OF INVENTION: COMPOSITION AND METHOD
FILE REFERENCE: 8900-0008.30
CURRENT APPLICATION NUMBER: US/09/865,159
CURRENT APPLICATION NUMBER: EARLIER APPLICATION NUMBER: EARLIER APPLICATION NUMBER: EARLIER PILING DATE: EARLIER FILING STEEL SOUTHONES: SOFTWARE: EASTSEQ for Windows Version 3.0
SEQ ID NO 20
LENGTH: 168
                                                                                                                                                                                                                                                                           OTHER INFORMATION: Description of Unknown Organism: Protein, human; OTHER INFORMATION: procalcitonin US-09-742-373-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 12, Application US/09865159
Fatent No. US20020098196A1
GENERAL INFORMATION:
APPLICANT: ITVIO, Randall T.
APPLICANT: Hodges, Robert S.
TITLE OF INVENTION: PSEUDOMONAS TREATMENT
TITLE OF INVENTION: COMPOSITION AND METHOD
FILE REFERENCE: 8900-0008.30
CURRENT APPLICATION NUMBER: US/09/865,159
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/329,884
                                                                                                                                                                                                                                                                                                                                                                                   Length 128;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24.9%; Score 94; DB 10; Length 168; larity 40.0%; Pred. No. 0.03; Conservative 8; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6 HHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 HHHHHGG-----GGEIEALKAEIEALKAEIEALKAEIEALKA 47
                                                                                                                                                                                                                                                                                                                                                                             25.3%; Score 95.5; DB 10; 30.8%; Pred. No. 0.017; tive 18; Mismatches 17;
PRIOR APPLICATION NUMBER: 10016278.9
PRIOR FLING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: 10027954.6
PRIOR FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 20, Application US/09865159
Patent No. US20020098196A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                59 REGSSLDSPRSKRCGNLS 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 30.89
Matches 24; Conservative
                                                                                                                                                                                                                           ORGANISM: Unknown Organism
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 20; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
US-09-865-159-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-865-159-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-865-159-20
```

ï

ä

Gaps

```
3 GSHHHHHHGSMASGDLK-----NKVAQLKRKV-----RSLKDKAAELKQEV 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Symersky, Jindrich
TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; OTHER INFORMATION: pT7H6 Trip-A-Fn-Apo Al-final - AmpR plasmid US-09-987-107-58
                                                                    Length 323;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 159;
                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION INMBER: US/09/813,820
FILING DATE: 22-Mar-2001
CLASSIFICATION: <Unknown>
                                                                    Score 93.5; DB 9;
Pred. No. 0.066;
8; Mismatches 27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 92; DB 10
Pred. No. 0.044;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AND METHODS OF USE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE: Arnold, White & Durkee STREET: P.O. Box 4433 CITY: Houston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/856,253
FILING DATE: «UNKNOMER: 08/856,253
ATTORNEY/AGENT INFORMATION:
NAME: KItchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: TAMK:193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear SEQUENCE DESCRIPTION: SEQ ID NO: 2:
                                                                                                                                                                                                                                                                      44 SRLENEIEDLKAKIGDLNNTSGIRRP 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                House-Pompeo, Karen
Sthanam, Narayana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
                                                                                                                                                                                                                                                                                                                                                                                            Sequence 2, Application US/09813820 Patent No. US20020102262Al GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: amino acid
STRANDEDNESS: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: (512) 474-7577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Patti, Joseph M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24.8%;
32.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Hook, Magnus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STATE: Texas
COUNTRY: U.S.
ZIP: 77210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 15; Conserva
                                                                       Query Match
Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                        US-09-813-820-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-813-820-2
                                                                                                                                                                                                 g
                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                              ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: Irvin, Randall T.
APPLICANT: Hodges, Robert S.
TITLE OF INVENTION: ESEUDOMONAS TREATMENT
TITLE OF INVENTION: COMPOSITION AND METHOD
TITLE OF INVENTION: COMPOSITION AND METHOD
CURRENT APPLICATION NUMBER: US/09/865,159
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/329,884
PRIOR AILING DATE: EARLIER FILING DATE: 1999-06-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24.9%; Score 94; DB 10; Length 175; llarity 40.0%; Pred. No. 0.032; Conservative 8; Mismatches 16; Indels
                                                                                                                                                                                             Score 94; DB 10; Length 169;
Pred. No. 0.031;
8; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                   6 HHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                                                                                                                                                                                                                                                                            4 HHHHHGG-----GGEIEALKAEIEALKAEIEALKAAEIEALKA 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
    PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-11
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 12
LENGTH: 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: GRAVERSEN, Jonas
APPLICANT: MCGSTRUP, SOTEN
TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
FILE REFERENCE: CRAVERSENIA
CURRENT FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/264,022
PRIOR PILING DATE: 2001-01-26
PRIOR FILING DATE: 2001-01-26
PRIOR FILING DATE: 2001-01-15
PRIOR PLICATION NUMBER: DK PA2001 00057
PRIOR PLICATION NUMBER: DK PA2000 01682
PRIOR PELICATION NUMBER: DK PA2000 01682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 16
LENGTH: 175
                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 16, Application US/09865159 Patent No. US20020098196A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-987-107-58
. Sequence 58, Application US/09987107
. Patent No. US/0020156007A1
. GENERAL INFORMATION:
                                                                                                             ; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-865-159-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; ORGANISM: Pseudomonas aeruginosa US-09-865-159-16
                                                                                                                                                                                               24.9%;
40.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQ ID NOS: 91
SOFWARE: Patentin version 3.1
SEQ ID NO 58
LENCTH: 323
                                                                                                                                                                                                                                       Matches 20; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Loca
Matches
                                                                                                                                                                                                                                                                                   ò
```

ö

Gaps

ö

```
Sequence 22, Application US/09865159
Patent No. US20020098196A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-865-159-14
                                                                                                                                                                                                                                                                                                                               SEQ ID NO 22
LENGTH: 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-865-159-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 14
LENGTH: 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65 G 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           58 G 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    65 G 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              58 G 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 GSHHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGD 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 GSHHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match

24.4%; Score 92; DB 9; Length 336; Best Local Similarity 40.4%; Pred. No. 0.094; Matches 23; Conservative 6; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24.4%; Score 92; DB 9; Length 337; llarity 40.4%; Pred. No. 0.095; Conservative 6; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; OTHER INFORMATION: T7 H6UbiFx Cys-Apo A-1 plasmid US-09-987-107-46
                                                                                              APPLICANT: GRAVERSEN, JODAS
APPLICANT: GRAVERSEN, JODAS
APPLICANT: GRAVERSEN, SOLE
ITLE OF INVENTION: APOLITOPROTEINS ANALOGUES
FILE REFERENCE: GRAVERSENIA
CURRENT APPLICATION NUMBER: US/09/987,107
CURRENT APPLICATION NUMBER: US 60/264,022
PRIOR FILING DATE: 2001-01-26
PRIOR FILING DATE: 2001-01-15
PRIOR PLILING DATE: 2001-01-15
PRIOR APPLICATION NUMBER: DK PA2000 01682
PRIOR PILING DATE: 2000-11-10
PRIOR FILING DATE: 2000-11-10
NUMBER: OF SEO ID NOS: 91
SOFTWARE: PATENTIN VETSION 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: GRAVERSEN, JODGS
APPLICANT: GRAVERSEN, JODGS
APPLICANT: GRAVERSEN, SOFEN
TITLE OF INVENTION: APOLIPOPROFIEINS ANALOGUES
FILE PEFERENCE: GRAVERSENIA
CURRENT APPLICATION NUMBER: US/09/987,107
CURRENT FILING DATE: 2001-11-13
PRIOR FILING DATE: 2001-01-26
PRIOR FILING DATE: 2001-01-26
PRIOR FILING DATE: 2001-01-15
PRIOR FILING DATE: 2001-01-15
PRIOR FILING DATE: 2000-11-10
SOFTWARE: PALCATION NUMBER: DK PA2000 01682
PRIOR FILING DATE: 2000-11-10
SOFTWARE: PALCATION NUMBER: DK PA2000 01682
SOFTWARE: PALCATION NUMBER: DK PA2000 01682
LENGTH: 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; FEATURE:
; OTHER INFORMATION: pT7 H6Ublfx Apo A-1 plasmid
US-09-987-107-44
                                         Sequence 44, Application US/09987107
Patent No. US20020156007A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; Sequence 46, Application US/09987107
; Patent No. US20020156007A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Artificial Sequence FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
RESULT 10
US-09-987-107-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-987-107-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

RESULT 12

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ñ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6 HHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDL-KAKIGDLNNTS 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 HHHHHGG------GGEVSALEKEVSALEKEVSALEKEVSALEKEVSALEKGGGEFARSE 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6 HHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDL-KAKIGDLNNTS 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 14, Application US/09865159
Patent No. US20020098196A1
Patent No. US20020098196A1
Fetent No. US20020098196A1
Fetent No. US20020098196A1

APPLICANT: Irvin, Randall T.
APPLICANT: Hodges, Robert S.
TITLE OF INVENTION: COMPOSTION NOT BEATMENT
FILE REFERENCE: 8900-0008.30

CURRENT APPLICATION NUMBER: US/09/865,159
CURRENT FILING DATE: EARLIER APPLICATION NUMBER: 09/329,884

PRIOR PLING DATE: EARLIER FILING DATE: 1999-06-11

NUMBER OF SEQ ID NOS: 22

SOFTWARE: FASTSEQ for Windows Version 3.0
APPLICANT: ITCHIN, Randall T.
APPLICANT: Hodges, Robert S.
TITLE OF INVENTION: COMPOSITION AND METHOD
FILLS REFERENCE: 8900-0008.30
CURRENT APPLICATION NUMBER: US/09/865,159
CURRENT APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/329,884
PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-06-11
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FASTER FILING DATE: EASTER FILING DATE: 1999-06-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 24.3%; Score 91.5; DB 10; Length 168; Best Local Similarity 34.4%; Pred. No. 0.051; Matches 21; Conservative 14; Mismatches 19; Indels 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 24.3%; Score 91.5; DB 10; Best Local Similarity 34.4%; Pred. No. 0.052; Matches 21; Conservative 14; Mismatches 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
US-10-007-805-551
Sequence 551, Application US/10007805
; Patent No. US20020150581A1
                                                                                                                                                                                                                                                                                                                                                                                      ; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-865-159-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Pseudomonas aeruginosa
```

```
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 HHHHHHGSMA-----SGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLEN---E 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 GSHHHHHHGSMAS------GDLKNKVAQLKRKV-----RSLKDKAAELKQEVSRLEN 48
                                          APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: ALLOCKET, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: McMedic, Thomas S.
APPLICANT: McWell, Patricta D.
APPLICANT: MWENTON: AND DIAGNOSIS OF BREAST CANCER
FILE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE FERENCE: 210121.470C10
CURRENT FILING DATE: 2001-12-07
CURRENT FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 593
SOFTWARRE: FastSEQ for Windows Version 4.0
SEQ ID NO 551
LENGTH: ALLOCATION CONTINUED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34; Indels 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24.3%; Score 91.5; DB 12; Length 324; 32.4%; Pred. No. 0.1; Live 11; Mismatches 24; Indels 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; OTHER INFORMATION: pT7H6 Trip-A-Apo A-1-del 43 - AmpR plasmid
US-09-987-107-50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 23.9%; Score 90; DB 9; Length 273; Best Local Similarity 29.2%; Pred. No. 0.12; Matches 26; Conservative 11; Mismatches 34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 50, Application US/09987107
Sequence 50, Application US/09987107
Sequence 50, Application US/09987107
Sequence 50, Application US/09987107
SEQUENT: GENERAL INFORMATION:
APPLICANT: MOESTRUP, Soren
TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
FILE REFERENCE: GRAVERSENIA
CURRENT APPLICATION NUMBER: US/09/987,107
CURRENT FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/264,022
PRIOR FILING DATE: 2001-01-26
PRIOR FILING DATE: 2001-01-15
PRIOR APPLICATION NUMBER: DK PA2000 01682
PRIOR FILING DATE: 2001-01-15
SPIOR FILING DATE: 2000-11-10
NUMBER OF SEQ ID NOS: 91
SEQ ID NO 50
LENGTH: 273
GENERAL INFORMATION:
APPLICANT: Jiang, Yuqiu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best_Local Similarity 32.4%
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          50 IEDLKAKIGDLNNT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |: || | :| |
IKILKEKNAELQMT 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
GRGANISM: Homo sapiens
US-10-007-805-551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
US-09-987-107-50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
```

```
49 ----EIEDLKAKIGDLNNTSGIRRPAAKL 73
                 62 QQALQTVSLKGSLKLLDNWDSVTSTFSKL 90
                          g
ò
```

Search completed: December 18, 2002, 16:39:47 Job time: 17.0223 secs

This Page Blank (USDIO)

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

December 18, 2002, 16:13:49; Search time 2.18649 Seconds (without alignments) 3253.588 Million cell updates/sec

Title: Perfect score: Sequence:

US-09-490-291-11 377 1 MRCSHHHHHHGSMASGDLKN.....AKIGDLNNTSGIRRPAAKLN 74

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283224 seqs, 96134422 residues Searched:

283224 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

SH	Description	chromosome segreda	somite Maf1 protei			transcription fact	Ig V-region-like B	conserved hypothet	P115 protein - Myc					transcription acti		w.	hypothetical profe								5-	myosin heavy chain	chromosome segreda	myosin heavy chain	hemagglutinin prec	
SUMMAKIES	ID	6970	JC7572	A70359	150167	149529	D39371	F75103	JQ0894	F75216	T29145	F90182	JC5610	A56235	S37894	683960	H69402	S73342	G71061	T41405	F72504	T18335	A84265	F64341	в60608	833068	A49464	A59287	HMIVE3	HMIVE4
	DB	7	~	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	~	~	~	7	7	~	Н	-
	Length	1186	315	259	155	323	321	880	979	281	1199	397	173	311	992	1188	212	166	270	396	409	269	647	102	103	527	1225	1940	565	265
æ	Query Match	25.1	23.6	23.3	22.8	22.7	22.5	22.5	22.5	22.4	22.3	21.9	21.8	21.8	21.8	21.8	21:6	21.5	21.5	21.5	21.5		21.4				•		21.1	21.1
	Score	94.5	89	88	86	85.5	82	82	82	84.5	84	82.5	82	82	82	œ	81.5	81	81	81	81	80.5	0	80	80	80	80	80	79.5	Ġ.
	Result No.	-	7	m	4	ഗ	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

hypothetical prote transforming prote hypothetical prote hypothetical prote hemagalutinin prec hemagalutinin prec hypothetical prote hypothetical prote plis homolog - Met myosin heavy chain myosin heavy chain heat shock protein	hypothetical protection protection from transcription fact retrovirus-related
T49167 TVFVAF F72343 E72361 HMIVE2 HMIVE2 HMIVE3 A44505 A4465 A4469 A440997	130430 F96587 I51171 I38587
0H00HH000H0	0000
318 318 309 4009 509 1199 11109 11133 848	778 1556 292 338
211.0 211.0 2011.0 20.1 200.7 200.7	20.6 20.6 20.4 20.4
79 79 78 78 78 78 78 78 78	77.5 77.5 77
0 H O H O H O H O H O H O H O H O H O H	4444 12642

## ALIGNMENTS

RESULT 1 G69708 Chromosome segregation SMC protein - Bacillus subtilis N.Alternate names: minichromosome stabilizing protein SMC
C;Species: Bacillus subtlins C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000 C;Accession: 669708; JC4819; PC4029 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Pabret, C.; Perrari, Nature 390, 249-256, 1997
A; Authors: Foulger, D.; Fritz, C.; Fujlta, M.; Fujlta, Y.; Fuma, S.; Galizzi, A.; Galiech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M. Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapldus, A.; Lardhor A; Airhora I anber, J.; Lazarevic, V. Teo, S. M. Lovino A. Iti H. Masuda G. Mai
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Dohl, T.M.; Portete Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadala, Y.; Sato, T.; Scanl A. Authora: Schleich, S.; Schroder, R. Schoffone, F. Sakinichi, T. Sakowaka, A. Sa
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida A.; Althors: Yoshidawa, H.F.; Zumatein, R.; Yoshikawa, H. Danchin, a
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili A;Reference number: A69580; MUID:98044033; PMID:9384377 A;Accession: G69708
A;Status: nucleic acid sequence not shown; translation not shown A;Molecule type: DNA
A; Kesidues: 1-1186 KNUN> A; Kesidues: 1-1186 KNUN> A; Cross-treferences: GB: 299112; GB: AL009126; NID: 92633902; PIDN: CAB13467.1; PID: 926339 A; Experimental source: strain 168 R; Oguro, A.; Kakeshita, H.; Takamatsu, H.; Nakamura, K.; Yamane, K.
Gene 172, 17-24, 1996 A;Title: The effect of Srb, a homologue of the mammalian SRP receptor alpha-subunit, A;Reference number: JC4819; MUID:96257247; PMID:8654983 A;Accession: JC4819
A;Molecule type: DNA A;Residues: 1-49,'G', 51-161,'G',163-174,'E',176-177,'G',179-191,'G',193-227,'P',229-2 74, P',476-493,'D',495-514,'D',516-541,'V',543-545,'P',547-585,'SKPLRGNSGPAFIISF',601 'E',741-1186, Octor
A; Cross. Teferences: DDBJ:D64116; NID:g1389548; PIDN:BAA10977.1; PID:g1237015 R; Guro, A.; Kakeshita, H.; Honda, K.; Takamatsu, H.; Nakamura, K.; Yamane, K. DNA Res. 2, 95-100, 1995
A;Title: srb: a Bacillus subtills gene encoding a homologue of the alpha-subunit of t. A;Reference number: JC4093; MUID:96093930; PMID:7584053 A;Accession: PC4029
A;Molecule type: DNA A;Residues: 1171-1186 <og2> A;Cross-references: DDBJ:D49781; NID:91237017; PIDN:BAA08615.1; PID:91237018 C;Genetics:</og2>
A;Gene: smc C;Superfamily: chromosome segregation protein SMC1

25.1%; Score 94.5; Query Match

DB 2; Length 1186;

```
RLENEIEDLK--AKIGDLNNTSGIRRPAAKL 73
         45
         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A Status: preliminary; nucleic acid sequence not shown; translation not shown A; Molecule type: DNA A; Molecule type: DNA A; Residues: 1-259 <AQF>
A; Residues: 1-259 <AQF>
A; Cross-treferences: GB: AEO00701; NID: 92983260; PIDN: AAC06860.1; PID: 92983266; GB: AEO0065 A; Experimental source: strain VF5
C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C;Comment: This protein, a member of large Maf protein subfamily, containing a basic-leu mportant roles in somitogenesis, and has specific functions as to the determination and C;Genetics: A;Gene: s-mafl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Accession: JC7572
R; Kajihara, M.; Kawauchi, S.; Kobayashi, M.; Ogino, H.; Takahashi, S.; Yasuda, K.
J. Blochem. 129, 139-146, 2001
A; Title: Isolation, characterization, and expression analysis of zebrafish large Mafs.
A; Reference number: JC7572; MUID: 21064923; PMID:11134968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hydrogenase expression/formation protein B - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 24-Sep-1999
C;Accession: A70359
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus. A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: A70359
                                                                                                                                                                                                                                                                                                                                                                                                                                         somite Maf1 protein - zebra fish
C;Species: Brachydanio rerio (zebra fish)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Superfamily: hydrogenase expression/formation protein hypB
      Pred. No. 2.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 23.6%; Score 89; DB Best Local Similarity 25.5%; Pred. No. 2.2; Matches 24; Conservative 12; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: DDBJ:AB006324
36.8%;
Best Local Similarity 36.89
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nature 392, 353-358, 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: mRNA
A; Residues: 1-315 <KAJ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                  RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 3
                                                                                                                                                  셤
                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
```

m

5

Gaps

10;

Length 979

```
submitted to the EMBL Data Library, July 1999
A; Reference number: A75001
A; Recession: F75216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A, Status: preliminary
A, Molecule type: DNA
A.Rolecule type: DNA
A.Rosidues: 1-281 < KAW>
A, Cross-references: GB: AJ248283; GB: AL096836; NID: g5457433; PIDN: CAB49181.1; PID: g545
A, Experimental source: strain Orsay
C, Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cypecies: Caenorhabditis elegans
Cypecies: Caenorhabditis elegans
Cypecies: 15-oct-1999 #sequence_revision 15-oct-1999 #text_change 15-oct-1999
Cyccossion: 129145
Cyccossion: T29145
AyDescription: The sequence of C. elegans cosmid F56A3.
AyDescription: The sequence of C. elegans cosmid F56A3.
AyAccession: T29145
AyAccession: T29145
AyAccession: T29145
AyAccession: T29145
AyAccession: T2919 GEI>
AyAccession: T2919 #text_chanslated from GB/EMBL/DDBJ
AyAccession: T29145
AyAccession: T29145
AyAccession: T29145
AyAccession: T29145
AyAccession: T29145
AyAccession: T2919 #sequence from GB/EMBL/DDBJ
AyAccession: T2919 #sequence from GB/EMBL/DBJ
AyAccession: T2919 #sequence from GB/EMBL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
                                                                                                                                                                                                                                                                                                                                                  17 DLKNKVAQLKRKVRSL-----KDKAAELKQEVSRLENEIEDLKAKIGDLNNTSGI--- 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Gene: CESP:F56A3.4
A;Map position: 1
A;Introns: 18/1; 80/3; 446/3; 475/1; 607/3; 679/1; 974/2; 1058/1; 1168/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hypothetical protein PAB2181 - Pyrococcus abyss1 (strain Orsay)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17 DLKNKVAQLKR-----KVRSLKDKAAELKQEVSRLENEIEDLKAKIG 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12 SMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNN
                                                                                                                                                                                                                              22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hypothetical protein F56A3.4 - Caenorhabditis elegans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB. 2;
       C;Superfamily: chromosome segregation protein SMC1 C;Keywords: nucleotide binding; P-loop F;32-39/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                  22.5%; Score 85; DB 2; 30.9%; Pred. No. 14; tive 15; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22.3%; Score 84; DB 2;
33.3%; Pred. No. 21;
iive 13; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22.4%; Score 84.5; D
40.8%; Pred. No. 4.6;
:1ve 10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 33.3%
Matches 17; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 22.4%
Best Local Similarity 40.8%
Matches 20; Conservative
                                                                                                                                                      Query Match 22.5%
Best Local Similarity 30.9%
Matches 21; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      383 KKSAAAAN 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    R; anonymous, Genoscope
                                                                                                                                                                                                                                                                                                                                                                                                                                                67 RRPAAKLN 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C; Accession: F75216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Gene: PAB2181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genetics
                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                     ð
                                     C;Accession: D39371
R;Miller, M.M.; Goto, R.; Young, S.; Chirivella, J.; Hawke, D.; Miyada, C.G.
Proc. Natl. Acad. Sci. U.S.A. 88, 4377-4381, 1991
A;Title: Immunoglobulin variable-region-like domains of diverse sequence within the majo A;Reference number: A39371; MUID:91239571; PMID:1903541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cybace: 20-Ang-1799; #Sequence_revision 20-Ang-1999; #Lext_Change 02-Feb-2001
Cybaces 20-Ang-1799; #Sequence_revision 20-Ang-1999; #Lext_Change 02-Feb-2001
Ryanonymous, Genoscope
Submitted to the EMBL Data Library, July 1999
A; Bescription: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struck A; Reference number: A75001
A; Accession: F75103
A; Accession: F75103
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 1-880 cKAMA
A; Cross-references: GB:AJ348286; GB:AL096836; NID:g5458366; PIDN:CAB50131.1; PID:g545864
A; Experimental source: strain Orsay
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Accession: J00894
R;Notarnicola, S.M.; McIntosh, M.A.; Wise, K.S.
Gene 97, 77-85, 1991
A;Title: A Mycoplasma hyorhinis protein with sequence similarities to nucleotide-binding A;Reference number: J00894; MUID:91138990; PMID:1825306
A;Accession: J00894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ñ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
C;Species: Gallus gallus (chicken)
C;Date: 20-Mar.1992 #sequence_revision 10-Apr-1992 #text_change 21-Jul-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C;Species: Mycoplasma hyorhinis
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 02-Feb-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7 HHHHGSMASGDLK------NKVAQLKRKVRSLKDKAAELKQEVSRLENEIED 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 conserved hypothetical protein PAB0812 - Pyrococcus abyssi (strain Orsay) C; Species: Pyrococcus abyssi
                                                                                                                                                                                                                                                                                          A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-321 «MIL»
A;Cross-references: GB:M61862; NID:g211259; PIDN:AAA48627.1; PID:g211261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A.Molecule type: DNA
A.Residues: 1-979 <NOT>
A.Cross-references: GB:M34956
A.Note: the authors translated the codon AAA for residue 956 as Leu
C.Comment: This protein is located in the cytoplasm.
C.Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 321;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22.5%; Score 85; DB 2; Length 880; 37.2%; Pred. No. 13; Live 14; Mismatches 13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              240 IKGKISELKIQVEKLKGRRKGLEEKIVQIERSIEEKRAKISEL 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 LKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22.5%; Score 85; DB 2;
26.8%; Pred. No. 4.8;
ive 18; Mismatches 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VEKHLKKI----GIRAPNLKLH 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LKAKIGDLNNTSGIRRPAAKLN 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P115 protein - Mycoplasma hyorhinis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 22.5%
Best Local Similarity 37.2%
Matches 16; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 22.5
Best Local Similarity 26.8
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Comment: This prot
C;Genetics:
A;Genetic code: SGC3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gene: PAB0812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
```

-:

Gaps

7;

Length 281;

ö

Gaps

ö

Length 1199;

4

Page

ä

16;

```
A) Residues: 1-766 -RAS>
A) Cross-references: EMBL:228072; NID:9486100; PIDN:CAA81909.1; PID:9486101; MIPS:YKLO
A) EXPOST-IMENTAL SOURCE: Strain $228C
B) POhl, T.M.; Pohl, F.M.
Submitted to the Protein Sequence Database, March 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              R;Rasmussen, S.W.
Yeast 10, 69-74, 1994
A;Title: Sequence of a 20.7 kb region of yeast chromosome XI includes the NUPl00 gene rp in addition to seven ORFS with weak or no significant similarity to known proteins A;Reference number: S44513
A;Accession: S44515
A;Status: nucleic acid sequence not shown; translation not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Cross-references: EMBL:X75780; NID:9433625; PIDN:CAA53402.1; PID:9433628
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N;Alternate names: hypothetical protein YKL352
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 29-Oct-1999
C;Accession: 33784; 33787; S39170; S44515
R;Rasmussen, S; von Wettstein, D.
Submitted to the Protein Sequence Database, March 1994
A;Reference number: S37872
                                                                                                                                                                                                                                                                                                                                                                                                                      A;Molecule type: DNA
A;Residues: 1-766 <RA2>
A;Cross-references: EMBL:X75780; NID:g433625; PIDN:CAA53402,1; PID:g433628
                                                                                                                                                                                                               Length 311;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hypothetical protein YKL072w - yeast (Saccharomyces cerevisiae)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2; Length 766
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25;
                                                                                                                                                                                                                                                                                          20;
A;Introns: #status absent
C;Superfamily: maf transforming protein; maf homology
C;Keywords: DNA binding; homodimer; leucine zipper
F;200-289/Domain: maf homology <WAF>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: DNA
A; Residues: 1-557 <POH>
A; Residues: 1-557 <POH>
A; Cross-references: EMBL: Z28072; MIPS:YKL072w
A; Experimental source: strain S288C
R; Rasmussen, S.W.
Bubmitted to the EMBL Data Library, November 1993
A; Reference number: S39168
A; Accession: S39170
                                                                                                                                                                                                                                                ; Pred. No. 8.1; 12; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 21.8%; Score 82; DB Best Local Similarity 26.7%; Pred. No. 20; Matches 20; Conservative 20; Mismatches
                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SGD:S0001555; MIPS:YKL072w
                                                                                                                                                                                                           21.8%; Score 82; 21.7%; Pred. No.
                                                                                                                                                                                                                                                                                                                                                                     5 HHHHHGSMAS------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17 DLKNKVAQLKRKVRSLKDK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              279 RDAYKLKCEKL-ASNGFR 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        50 IEDLKAKIGDLNNTSGIR 67
                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Reference number: $37897
A; Accession: $37897
                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 30; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1-766 <RA3>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: DNA A; Residues: 1-766 <R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Gene: SGD:STB6
A;Cross-references:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Accession: S37894
                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                      Chan-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cippenia: 1 sea squir. Insacrymenta rotecal,
Cippenia: 1 sea squir. Insacrymenta rotecal,
Cipate: 23-Sep-1997 #sequence_revision 17-oct-1997 #text_change 20-Jun-2000.
Cincession: JC5610
Cincession: JC5610
Cincession: J22, 374-380, 1997
A;Title: Primary structure of troponin I isoforms from the ascidian Halocynthia roretzi.
A;Reference number: JC5610; MUID: 98021076; PMID: 9378716
A;Reference number: JC5610
A;Molecule type: mRNA
A;Residues: 1-173 <VUA>
A;Residues: 1-173 <VUA>
A;Residues: 1-173 <VUA>
A;Cross-references: DDBJ-AB001685; NID: 91888344; PIDN: BAA19425.1; PID: 91888345
                                                                                                                                                                                                                                                         Cyncusaries so to the state of 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                that can associate with Maf and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A) Experimental source: adult Halocynthia roretzi
C;Comment: This protein binds to actin, and inhibits the interaction between actin and
C;Superfamily: troponin I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lianscription activator MafB - chicken Cispecies: Gallus gallus (chicken) Cispecies: Gallus gallus (chicken) Cispecies: Gallus gallus (chicken) Cispecies: O3-Oct-1995 #sequence_revision O3-Oct-1995 #text_change 20-Jun-2000 Cisacession: A56235 Rstaoka, K.; Fujiwara, K.T.; Noda, M.; Nishizawa, M. Mol. Cell. Biol. 14, 7581-7591, 1994 A; Title: MafB, a new Maf family transcription activator that can associate w A; Reference number: A56235; MUID:95021288; PMID:7935473
                                                                                                                hypothetical protein SSO0390 [imported] - Sulfolobus solfataricus:
C;Species: Sulfolbus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: F90182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11 GSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNNTSGIRRP 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <KAT>
A;Cross-references: GB:D28600; NID:g516723; PIDN:BAA05938.1; PID:g516724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18 LKNKVAQLKRKVRSLKDKAAELKQ-----EVSRLENEIEDLKAKIGDLNNTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 21.8%; Score 82; DB 2; Length 173; Best Local Similarity 30.5%; Pred. No. 4.5; Matches 18; Conservative 16; Mismatches 23; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ch 21.9%; Score 82.5; Di
1 Similarity 34.5%; Pred. No. 9.4;
19; Conservative 13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     troponin I - sea squirt (Halocynthia roretzi)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Gene: $500390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

ä

Gaps

10;

```
C; Species: Bacillus halodurans (c; Dec-2000 #text_change 17-May-2002 (c; Decession: 683960 (c; Takaki, Y.; Maeno, G.; Sasaki, R.; Masul, N.; Fuji, F.; Hira Nucleic Acids Res. 28, 4317-4331, 2000 (c; Decession: G83960 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      chromosome segregation SMC protein smc [imported] - Bacillus halodurans (strain C-125)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1;
                                                            Query Match 21.8%; Score 82; DB 2; Length 1188; Best Local Similarity 35.0%; Pred. No. 31; Matches 21; Conservative 12; Mismatches 17; Indels 10; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 GSMASGDLKNK------VAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDL 60
2 RGSHHHH-HHGSMAS------GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIE 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Gene: smc
C;Superfamily: chromosome segregation protein SMCl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: December 18, 2002, 16:22:25 Job time: 4.18649 secs
                                                                                                                                                                                                                                                                                                                                          |:: |: : ||: |
680 DVRKKLLEQNNSKDI 694
                                                                                                                                                                                                                                                  52 DLKAKIGDLNNTSGI 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                         g
                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                              a
```

This Page Blank (uspto)

```
GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2002, 16:06:04; Search time 1.13584 Seconds

(without alignments)

2702.183 Million cell updates/sec
```

Title: US-09-490-291-11
Perfect score: 377
Sequence: .1 MRGSHHHHHHGSMASGDLKN.....AKIGDLNNTSGIRRPAAKLN 74

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

្ន	<b>43</b> 4	1	ಥ	ø	ø	ø	•	Ø	ದ
onchocerca influenza	influenza	•	~	influenza	influenza	influenza	influenza	influenza	influenza
P21249 P03441	P12582	P12584	P12585	P12587	P12588	P43258	P43259	P12589	P11134
ANT1_ONCVO HEMA_IABAN	HEMA_IADH1	HEMA_IADH3	HEMA_IADH4	HEMA_IADH6	HEMA_IADH7	HEMA_IADHL	HEMA_IADHM	HEMA_IAME6	HEMA_IA2H3
	Ч-	Η.	Н	Н	Н	Н	Н	-	-
2022	550	550	550	550	550	550	550	550	550
19.6 19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5
73.5	73.5	73.5	73.5	73.5	73.5	73.5	73.5	73.5	73.5
34 35	36	38	38	40	41	42	43	44	45

## ALIGNMENTS

```
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                      P54842;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                        RESULT
          FFFS
                                                                                                                                                         à
                                                                                                                                                                                      요
                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                     염
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                             CHARACTERIZATION.
MEDLINE-98367134; PubMed-9701812;
MOTIVE S., Tsujikawa E., Hassan A.K., Asai K., Kodama T.,
Ogasawara N.;
"A Bacillus subtilis gene-encoding protein homologous to eukaryotic
SMC motor protein is necessary for chromosome partition.";
Mol. Microbiol. 29:179-187(1998).
[5]
Kolaracterization.
MEDLINE-98241483; PubMed-9573042;
Britton R.A., Lin D.C., Grossman A.D.;
Characterization of a prokaryotic SMC protein involved in chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -> SKPLRGNSGPAFIISF (IN
                                  "srb: a Bacillus subtilis gene encoding a homologue of the alphasubunit of the mammalian signal recognition particle receptor."; DNA Res. 2:95-100(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E -> G (IN REF. 1).
E -> G (IN REF. 1).
E -> G (IN REF. 1).
E -> G (IN REF. 1).
E -> G (IN REF. 1).
A -> P (IN REF. 1).
B -> D (IN REF. 1).
C -> G (IN REF. 1).
C -> D (IN REF. 1).
C -> C (IN REF. 1).
C -> D (IN REF. 1).
C -> D
Oguro A., Kakeshita H., Honda K., Takamatsu H., Nakamura K.,
Yamane K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NRSDYRGLKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A -> S (IN REF. 1).
S -> T (IN REF. 1).
E -> G (IN REF. 1).
A -> S (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Subtilist; BG11338; smc.
InterPro; IPR003439; ABC_transportr.
InterPro; IPR003405; SMC_C.
InterPro; IPR003405; SMC_N.
Pfam; PF02463; SMC_N; 1.
Pfam; PF02483; SMC_C; 1.
ATP-binding; Coiled coil; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REF. 1).
TVLITEDLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D64116; BAA10977.1; -.
EMBL; Z99112; CAB13467.1; -.
EMBL; D49781; BAA08615.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           631
664
676
680
694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NP_BIND
DOMAIN
DOMAIN
DOMAIN
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
```

```
;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN-Wister;
STRAIN-Wister;
MEDLINE-97190228; Pubmed-9038383;
Sakai M., Imaki J., Yoshida K., Ogata A., Matsushima-Hibaya Y.,
Kuboki Y., Nishizawa M., Nishi S.;
Rat maf related genes: specific expression in chondrocytes, lens and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Métazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                        11 GSMASGDLKNKVAQLKRKVRSLKD------KAAELKQEVSRLENEIEDLKAKIGDL 60
                                                                                                                                                                          21; Indels 11; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .- 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    tion; Repressor; DNA-binding; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Transcription factor MafB (V-maf musculoaponeurotic fibrosarcoma
oncogene homolog B) (Transcription factor MAF1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   85;
                                                                                                                       Query Match 25.1%; Score 94.5; DB 1; Length 1186; Best Local Similarity 36.8%; Pred. No. 0.79; Matches 25; Conservative 11; Mismatches 21; Indels 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ch 22.9%; Score 86.5; DB 1; Length 323; 1 Similarity 23.1%; Pred. No. 1; 34; Conservative 9; Mismatches 19; Indels 85
701 701 K -> Q (IN REF. 1).
726 726 L -> V (IN REF. 1).
738 740 LQV -> PQF (IN REF. 1).
1186 Aa; 135510 MW; 0163227AAF2BCA2B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --GSMAS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLY-ALA.
6E386340D1F840A5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LEUCINE-ZIPPER.
POLY-HIS.
POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 323 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASIC MOTIF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPRO04826; TF_Maf.
InterPro; IPRO04827; TF_DZIP.
Pfam; PF03131; bZIP_Maf; 1.
SMART; SM00338; BRLZ: 1.
Transcription regulation; Repred DNA_BIND 238 264 BA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                35792 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; U56241; AAB50062.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 нининн ------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             323 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                     | |:|
718 RETGEGLR 725
                                                                                                                                                                                                                                                                                                                      61 NNT-SGIR 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MAFB OR MAF1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA_BIND
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MAFB_RAT
                                                                                                                          Query Match
```

3

```
MAFB_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.

IZARD J., Chartier M., Baty D.;
Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.

-!- FUNCTION: THIS COLICIN IS A CHANNEL-FORMING COLICIN. THIS CLASS OF TRANSMEMBRANE TOXINS DEPOLARIZE THE CYTOPLASMIC MEMBRANE, LEADING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pilsl H., Braun V.; "Strong function-related homology between the pore-forming colicins K and 5.";
                                                                                               :|| : | :| :| :| :| | :|| :| | :|| :| | :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|
162 HHHHHHQASPPPSSAASPAQQLPTSHPGPGPHAAAAATAAGSNGSVEDRFSDDQLVSMSV 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TO DISSIPATION OF CELLULAR ENERGY.
--- FUNCTION: COLICINS ARE POLYBEPTIDE TOXINS PRODUCED BY AND ACTIVE AGAINST, ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.
--- SIMILARITY: BELONGS TO THE CHANNEL FORMING COLICIN FAMILY.
                                                               40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----GDLKNKVAQLKRKVRSLKDK------AAELK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLN 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15; Indels 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22.8%; Score 86; DB 1; Length 548;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -> G (IN REF. 2).
2E67689D088CCF86 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRINTS; PR00280; CHANLCOLICIN.
Probom; PD002657; Channel_colicin; 1.
PR05TE; PS00276; THANNEL_COLICIN; 1.
Antibiotic; Bacteriocin; Plasmid; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                         15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32.1%; Pred. No. 1.9; ive 13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP; P04480; 1COL.
InterPro; IPR000293; Channel_colicin.
Pfam; PF01024; Colicin; 1.
                                                                                                                                                                                                                    282 QEVSRLARERDAYKVKCEKLAN-SGFR 307
                                                                                                                                                                                 QEVSRLENEIEDLKAKIGDLNNTSGIR 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-96074330; PubMed-7592493;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteriol. 177:6973-6977(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           59661 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; X87834; CAA61099.1; -. EMBL; U27452; AAB41288.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity, 32.1%
les 18; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           155 1
548 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Plasmid ColK-K235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Escherichia.
NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Colicin K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                         CEAK_ECOLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                               16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5
                                                                                                                                                                                                                                                                                                                                            RESULT
   셤
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                               à
                                                                                                                  셤
                                                                                                                                                                                    δ
                                                                                                                                                                                                                                             g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACUBOLE FROM N.A.

WEDLINE-21638749; PubMed=11780052;

A Deloukas P., Matthews L.H., Ashurst J., Bubton J., Gilbert J.G.R.,

A Jones M., Stavindes G., Almeida J.P., Babbage A.K., Bagguley C.L.,

Balley J., Barlow K.F., Bates K.N., Beare D.M.,

Balley J., Barrill W.D. Balter S.E., Bridgeman A.M., Brown A.J.,

Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA Chapman D.V., Griffiths R.E., Connor R.E., Corby N.R.,

Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,

Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

Hummond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,

Lehvaslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA M.D., Matry D., Moconacolle L.J., Naclay R., Mokurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

Rather E.B.J.C.T., Prathalingam S.R., Plumb R.W., Rammasy H.,

Rhillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Rammasy H.,

Swann R.M., Sycamore N., Taylor R., Thomas D.W., Thorpe A.,

Swann R.M., Sycamore N., Taylor R., Thomas D.W., Thorpe A.,

Swann R.W., Sycamore P., Wallie, D.L., Williams L., Williams S.A.,

Wiltehead S.L., Whittaker P., Wallie, D.L., Williams L., Williams S.A.,

Wiltehead S.L., Whittaker P., Walle, D. L., Williams L., Williams B. A.,

Williang L., Wray P.W., Hubbard T., Durblin R.W., Bentley D.R.,

Ran Miltehead S.L., Wilterker P., Willie, D. L., Wallie, Beck S.,

Ran Milling R., Wiltehead S.L., Wiltehead S.L.,

Ran William B. Williams B. Williams B. W., Ranker B. Williams B. Williams B. W., Ranker B. Williams B. W., Ranker B. W., Ranker B. W., Ranker B. Williams B. Williams B. W., Ranker B. W., Ranker B. W., Ranker B. W.
                                                                                                                                                                                                                                                                                                                                                                                                                                     a candidate tumor suppressor gene in myeloid leukemias.";
Genomics 59:275-281(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "The DNA sequence and comparative analysis of human chromosome 20.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ... *14:005-071(2001).
FUNCTION: Plays a pivotal role in regulating lineage-specific
                                                                                                                                                                                                                                    Cranlata; Vertebrata; Euteleóstom1;
Catarrhin1; Hominidae; Homo.
          09Y5Q3; Q9H1F1;
15-JUN-2002 (Rel. 41, Created)
15-JUN-2002 (Rel. 41, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Transcription factor MafB (V-maf musculoaponeurotic fibrosarcoma
oncogene homolog B).
                                                                                                                                                                                                                                                                                                                                                               TISSUE-Bone marrow;
MEDLINE-99375320; PubMed-10444328;
Wang P.W., Elsenbart J.D., Cordes S.P., Barsh G.S., Stoffel M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hematopolesis by repressing ETSI-mediated transcription of erythroid-specific genes in myeloid cells (By similarity).
-!- SUBCELLURR LOCATION: Nuclear (By similarity).
-!- SIMILARITY: BELONGS TO THE BZIP FAMILY. NAF SUBFAMILY.
  323 AA
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AF134157; AAD30106.1; -.
EMBL; AL035665; CAB75863.1; -.
                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genew; HGNC:6408; MAFB.
InterPro; IPR004827; TF_bZIP.
  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nature 414:865-871(2001)
                                                                                                                                                                                                         Homo saptens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                       FROM N.A.
                                                                                                                                                                                                                                                                                    NCBI_TaxID=9606;
MAFB_HUMAN
                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Rogers
```

```
ò
                                                                                                                                                                                                        g
                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                    4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                 01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Transcription factor MafB (V-maf musculoaponeurotic fibrosarcoma oncogene homolog B) (Transcription factor MAFI) (Segmentation protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY, GUT, LUNG, AND BRAIN. DEVELOPMENTAL STAGE: DETECTABLE AT 8.0 DPC (ONE SOMITE) AS A BAND IN THE CAUDAL HINDBRAIN, AND BY 8.5 DPC (SIX TO EIGHT SOMITES), THE HIGH LEVEL DOMAIN EXHIBITS A SHARP ROSTRAL EDGE COINCIDENT WITH THE R4/R5 BOUNDARY AND A DIFFUSE CAUDAL EDGE LOCATED MIDWAY THROUGH R6.
                                                                                                                                                                                        162 HHHHHQASPPPSSAASPAQQLPTSHPGPGPHATASATAAGGNGSVEDRFSDDQLVSMSV 221
                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   regulation; Repressor; DNA-binding; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus
                                                                                                                                                    85;
                                                                                                                          22.7%; Score 85.5; DB 1; Length 323; larity 23.1%; Pred. No. 1.2; Conservative 9; Mismatches 19; Indels 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY: BELONGS TO THE BZIP FAMILY. MAF SUBFAMILY.
                                                                                                                                                  19; Indels
                                                                                                         A0F3C09F8936CB16 CRC64;
                                                                                   A -> V (IN REF. 1).
Q -> H (IN REF. 1).
                                                                                                                                                                    нннин------
                                        BASIC MOTIF.
LEUCINE-ZIPPER.
                                                                                                                                                                                                                                                                                                                                   323 AA.
                                                                                                                                        Pred. No. 1.2;
9; Mismatches
                                                              POLY-HIS. POLY-HIS.
                                                                                                                                                                                                                                                                           282 QEVSRLARERDAYKVKCEKLAN-SGFR 307
                                                                                                                                                                                                               ----GDLKNKVAQLKRKVRSLKDK---
                                                                                                                                                                                                                                                      41 QEVSRLENEIEDLKAKIGDLNNTSGIR 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUBCELLULAR LOCATION: Nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-95094266; PubMed-8001130;
                                                                                           241 0
35792 MW;
InterPro; IPR004826; TF_Maf.
         Pfam; PF03131; bZIP_Maf; 1.
SMART; SM00338; BRLZ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; L36435; AAA65689.1; -.
                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                 287
143
167
52
                                                                                                                                                                                                                                                                                                                                                                                                      KR) (Kreisler).
MAFB OR MAF1 OR KRML.
Mus musculus (Mouse).
                                                                                           241 2
323 AA;
                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSFAC; T01439;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=10090;
                               Transcription
                                                                                                                                               34;
                                                                                                                                                                                                                                                                                                                      MAFB_MOUSE
ID MAFB_MOUSE
                                       DNA_BIND
DOMAIN
DOMAIN
                                                                                 CONFLICT
CONFLICT
SEQUENCE
                                                                                                                          Query Match
                                                                                                                                       Local
                                                                                                                                                                                                                                                                                                                                         P54841;
                                                                       DOMAIN
                                                                                                                                               Matches
q
                                                                                                                                                                                                                                 g
                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                         g
```

```
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                     162 HHHHHHQASPPPSSAASPAQQLPTSHPGPGPHATAAATAAGGNGSVEDRFSDDQLVSMSV 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           222 RELNRHLRGFTKDEVIRLKOKRRTLKNRGYAQSCRYKRVQOKHHLENEKTOLIQOVEOLK 281
                                                                                                                                                                                                                                                                                                                                                                      -- 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pyrococcus abyssi genome sequence: insights into archaeal chromosome
                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GDLKNKVAQLKRKVRSLKDK--------AAELK 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaee; Pyrococcus.
                                                                                                                                                                                                                N->S: LOSS OF TRANSCRIPTIONAL ACTIVITY. D77AE07ABD9C2AD2 CRC64;
                                                                                                             DNA-binding; Nuclear protein.
                                                                                                                                                                                                                                                                                                                            85;
                                                                                                                                                                                                                                                                                  DB 1; Length 323;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -1- SUBUNIT: Forms a complex with mrell (By similarity).
-1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
                                                                                                                                                                                                                                                                                                                          19; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
DNA double-strand break repair rad50 ArPase.
                                                                                                       regulation; Repressor; DNA-bin
18 264 BASIC MOTIF.
16 287 LEUCINE-ZIPPER.
                                                                                                                                                                                                                                                                                                      Pred. No. 1.2;
9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       880 AA.
                                                                                                                                                                                                                                                                                  Score 85.5;
                                                                                                                                                                                                                                                                                                                                                                 5 ннннн------СЅМАЅ------
                                                                                                                                                                                          POLY-HIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Interpro; IPR003593; AAA_Arpase.
Interpro; IPR003439; ABC_transportr.
Interpro; IPR003395; SMC_N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF02463; SMC_N; 1. ProDom; PD000006; ABC_transportr; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    41 QEVSRLENEIEDLKAKIGDLNNTSGIR 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                         35809 MW;
MGD; MGI:104555; Mafb.
InterPro; IPR004826; TF_Maf.
InterPro; IPR004827; TF_bZIP.
Pfam: PF03131; bZIP_Maf; 1.
SMART; SM00338; BRLZ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AJ248286; CAB50131.1; -
                                                                                                                                                                                                                                                                              22.7%;
                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                         131
158
148
248
323 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-GE5 / Orsay;
                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pyrococcus abyssi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=29292;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RAD50 OR PAB0812
                                                                                                         Transcription
DNA_BIND 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              similarity
                                                                                                                                                                                                                                                                              Query ...
Best Local Simir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RA50_PYRAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Heilig R.;
                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                MUTAGEN
                                                                                                                                                                         DOMAIN
                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RA50_PYRAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
```

us-09-490-291-11.rsp

383 KKSAAAAN 390

a

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics. Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -1- SUBCELLULAR LOCATION: CYLOPIASMIC.
-1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COLLED-COLL REGIONS
FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.
-1- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYCOPLASMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17 DLKNKVAQLKRKVRSL-----KDKAAELKQEVSRLENEIEDLKAKIGDLNNTSGI--- 66
                           DNA repair; Hydrolase; ATP-binding; Colled coil; Complete proteome. Mp.BIND 30 37 ATP (BY SIMILARITY). DOMAIN 144 745 COLLED COIL (POTENTAL). SEQUENCE 880 AA; 103970 MW; FDB177EC7E026479 CRC64;
                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mycoplasma hyorhinis.
Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE-91138990; Pubmed-1825306;
Notarnicola S.M., McIntosh M.A., Wise K.S.;
"A Mycoplasma hyorhinis protein with sequence similarities nucleotide-binding enzymes.";
                                                                                                                                                                           Score 85; DB 1; Length 880; Pred. No. 3.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22.5%; Score 85; DB 1; Length 979; 30.9%; Pred. No. 4.2;
                                                                                                                                                              22.5%; Score bJ, 37.2%; Pred. No. 3.7; **ive 14; Mismatches 13; Indels **ive 14; Mismatches 14; Mismatches 14; Mismatches 15; Indels **ive 14; Mismatches 14; Mismatches 14; Mismatches 15; Indels **ive 14; Mismatches 14; Mismatches 15; Indels **ive 14; Mismatches 15; Mismat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             884 912 ALA/ASP-RICH (DA-BOX).
979 AA; 110566 MW; 30D51C56B56280F4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATP (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                             240 IKGKISELKIQVEKLKGRKKGLEEKIVQIERSIEEKKAKISEL 282
                                                                                                                                                                                                                                                                                            18 LKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1995 (Rel. 32, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Pred. No. 4.2;
15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR003439; ABC_transportr.
InterPro; IPR003489; GTP-bindding_dom.
InterPro; IPR003405; SMC_C.
InterPro; IPR003395; SMC_N.
Pfam; PF02463; SMC_N; 1.
Pfam; PF02463; SMC_C; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF02483; SMC_C; 1.
TIGRFAMS; TIGR00650; MG442; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M34956; AAA25423.1; -.
                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Coiled coil.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
SMART; SM00382; AAA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    224
400
821
912
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; JQ0894; JQ0894.
                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gene 97:77-85(1991)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-2100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        67 RRPAAKLN 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P115 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP-binding;
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P115_MYCHR
P41508;
                                                                                                                                                                                                                                     16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
NA KW
FT
SO
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              а
```

ò

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announco/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |||: | :|: | ::: | ::: |
620 RGSNEFHWEYGNWQNKQQILESEKKDMKSLSSKLKYNVRILDRRVRDVEASVDHFDRKLE 679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 RGSHHHH-HHGSMAS------GDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIE 51
                                                                                                                                                                                                                                                                                 "Sequence of a 20.7 kb region of yeast chromosome XI includes the
                                                                                                                                                                                                                                                                                                  NUP100 gene, an open reading frame (ORF) possibly representing a nucleoside diphosphate kinase gene, tRNAs for His, Val and Trp in addition to seven ORFs with weak or no significant similarity to
                                                                                                                                       Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungl; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STB1-STB5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21.8%; Score 82; DB 1; Length 766; 26.7%; Pred. No. 5.8; ive 20; Mismatches 25; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25; Indels
                                                                                                                                                                                                                                                                                                                                                                                                              Pohl T.M., Pohl F.M.;
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; X75780; CAA53402.1; -.
EMBL; Z28072; CAA81909.1; -.
PIR; S37894; S37894.
PIR; S39170; S39170.
PIR; S44515; S44515.
SGD; S0001555; YKL072W.
SEQUENCE 766 AA; 88835 MW; 9BCBEA2EE03A9AF1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Rel. 29, Created)
(Rel. 29, Last sequence update)
(Rel. 40, Last annotation update)
                                766 AA
                                                                                                                                                                                                                                                  MEDLINE-94378724; PubMed-8091863;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-98055153; PubMed-9393435;
                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 1-557 FROM N.A.
                                                                                                            STB6 protein.
STB6 OR YKL072W OR YKL352.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20; Conservative
                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                 known proteins.";
Yeast 10:S69-S74(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |::|: |: |
DVRKKLLEQNNSKDI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    52 DLKAKIGDLNNTSGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
Matches 20; Conserv
                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARACTERIZATION.
                                                                                                                                                                                         NCBI_TaxID-4932;
                                                                                                                                                                                                                                                                    Rasmussen S.W.;
                                                          01-JUN-1994 (
01-JUN-1994 (
16-OCT-2001 (
                                                                                                                                                                                                                                     STRAIN-S288c;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 9
RA50_PYRFU
                STB6_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   680
RESULT
                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

ဖ

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -i- FUNCTION: Involved in DNA double-strand break repair (DSBR). The rad50/mrell complex possesses single-strand endonuclease activity and ATP-dependent double-strand-specific exonuclease activity. Rad50 provides an ATP-dependent control of mrell by unwinding and/or repositioning DNA ends into the mrell active site.
-i- SUBGNIT: Forms a complex with mrell.
                                                                                                                                                     Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
                                                                                                                                                                                                                                                                                                                         "Mrell and Rad50 from Pyrococcus furiosus: cloning and blochemical characterization reveal an evolutionarily conserved multiprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 1-149.
STRAIN-VC1 / DSM 3638 / ATCC 43587 / JCM 8422;
MEDLINE-20348813; PubMed-10892749;
HOPfiner K.-P., Rarcher A., Shin D.S., Craig L., Arthur L.M.,
Carney J.P., Tainer J.A.;
"Structural biology of Rad50 ATPase: ATP-driven conformational
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA repair; Hydrolase; ATP-binding; Coiled coil; 3D-structure;
                                                                                                                                                                                                                                                                                    Hopfner K.-P., Karcher A., Shin D., Fairley C., Tainer J.A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   control in DNA double-strand break repair and the ABC-ATPase
                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-VC1 / DSM 3638 / ATCC 43587 / JCM 8422;
Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
"The complete sequence of the Pyrococcus furiosus genome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21.8%; Score 82; DB 1; Length 882; 29.1%; Pred. No. 6.7; Live 15; Mismatches 24; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          148 744 COILED COIL (POTENTIAL).
882 AA; 103839 MW; 3ADCBD250382A99E CRC64;
                                                                                                                                                                                                                                              STRAIN-VC1 / DSM 3638 / ATCC 43587 / JCM 8422;
MEDLINE-20485553; Pubmed=11029422;
                               16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
DNA double strand break repair rad50 ArPase.
882 AA.
                                                                                                                                                                                                                             SEQUENCE FROM N.A., AND CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR003439; ABC_transportr.
InterPro; IPR003395; SMC_N.
 PRT;
                                                                                                                                                                                                                                                                                                                                                                                 Bacteriol. 182:6036-6041(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AE010225; AAL81291.1; -.
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     superfamily.";
Cell 101:789-800(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF02463; SMC_N; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PDB; 1F2T; 02-AUG-00.
PDB; 1F2U; 02-AUG-00.
                                                                                                                                  Pyrococcus furiosus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                       Pyrococcus.
NCBI_TaxID=2261;
                                                                                                              RAD50 OR PF1167
                                                                                                                                                                                                                                                                                                           Carney J.P.;
RA50_PYRFU
P58301;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NP BIND
                                                                                                                                                                                                                                                                                                                                                                 machine
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                              'Complete sequence analysis of the genome of the bacterium Mycoplasma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kawarabayasi Y., Hino Y., Horikwa H., Yamazaki S., Haikawa Y., Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H., Takaniya M., Masuda S., Funahashi T., Tanaka T., Kubota K., Namazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K., Nakamura Y., Nomura N., Sako Y., Kikuchi H., "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix Kl.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15-JUN-2002 (Rel. 41, Created)
15-JUN-2002 (Rel. 41, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Proteascome-activating nucleotidase (Proteascome regulatory subunit).
PAN OR APE2012.
                                                                                                                                                                                 Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
0
                                                                                                                                                                                                                                                                                            Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21.5%; Score 81; DB 1; Length 166; 37.2%; Pred. No. 1.5; ive 15; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AE000003; AAB95664.1; -.
InterPro; IPR002862; DUF16.
Pfam; PF01519; DUF16; I
Hypothetical protein; Complete proteome.
SEQUENCE 166 AA; 19518 MW; BE44F5377B2FA709 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SI IENKVDKLEVKVDKLEEKVDKLEEKVDKLEEKVDKLEAKVDKL 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18 LKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDL 60
                                                                                                                                                                                                                                                                                                                                                                 Nucleic Acids Res. 24:4420-4449(1996).
-1- SIMILARITY: BELONGS TO THE UPF0134 FAMILY.
                                                                                         16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                       166 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           409 AA
                                                                                                                               protein MPN138 (E07_orf166).
                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-99310339; PubMed-10382966;
                                                                                                                                                                                                                                                     STRAIN=ATCC 29342 / M129;
MEDLINE=97105885; PubMed=8948633;
                                                                          (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Desulfurococcaceae; Aeropyrum.
NCBI_TaxID=56636;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 37.2% tes 16; Conservative
                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                               Mycoplasma pneumoniae.
                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                      NCBI_TaxID=2104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Aeropyrum pernix.
                                                                                                                                              MPN138 OR MP016
                                                                                                                               Hypothetical
                                                                                                                                                                                                                                                                                                                 Herrmann R.;
                                                                       16-OCT-2001
                                                                                                                                                                                                                                                                                                                                                    pneumoniae.'
                                     YB38_MYCPN
P75260;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PSMR AERPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-K1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
RESULT 10
YB38_MYCPN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PSMR_AERPE
                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
```

ö

Gaps

ö

Local Similarity 29.1

Best Loca Matches

ð

11 GSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNNTSG 65

```
Query Match
                                                                                                                                                                                                                                                                                     Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMC1_YEAST
          SO FW DR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                      between the Swiss Institute of Bioinformatics and the EMBL outstation—
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-96337999; PubMed-868087;
Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
Karlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
Utterback T.K., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
"Complete genome sequence of the methanogenic archaeon, Methanococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
       of
FUNCTION: Required for the ATP- or CTP-dependent degradation proteins, but not small peptides, by the 20S proteasome (By
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 GSHHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDL 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21.5%; Score 81; DB 1; Length 409; llarity 37.3%; Pred. No. 3.7; Conservative 11; Mismatches 19; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Archaea; Euryarchaeota; Methanococci; Methanococcales; Methanocaldococcaceae; Methanocaldococcus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            rP (POTENTIAL).
72F1274F0AB885DD CRC64;
                                                                                               SUBUNIT: Homohexamer (Potential).
SUBCELLULAR LOCATION: Cytoplasmic (By similarity)
SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AP000063; BAA81022.1; -.
INTERPO; IPR003593; AAA_ATPase_centr.
Féan; PF00104; AAA; 1.
SMART; SM00382; AAA; 1.
FNSITE; PS00674; AAA; 7.
TIGREAMS; TIGR01242; 265945; 1.
PPOTEASORNE; APP-binding; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-CCT-2001 (Rel. 40, Last annotation update)
Hypothetical protein MJ0334.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 1 J | SEQUENCE FROM N.A. | SEQUENCE FROM N.A. | SEQUENCE A 3067; | S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        409 AA; 44683 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Science 273:1058-1073(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; U67487; AAB98322.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Methanococcus jannaschii.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
nes 19; Conserv
                                                                              similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=2190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Y334_METJA
Q57780;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    annaschii
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X334_METJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11D
DD111D
DD111D
DD111D
DD11D
```

```
ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIANE-94103320; PubMed-8276886; Strunnikov A.V., Larionov V.L., Koshland D.; Strunnikov A.V., Larionov V.L., Koshland D.; "SMC1: an essential yeast.gene encoding a putative head-rod-tail protein is required for nuclear division and defines a new ubiquitous protein family."; J. Cell Biol. 123:1635-1648(1993).
                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chromosome segregation protein SMC1 (DA-box protein SMC1).
SMC1 OR CHLIO OR YFL008W.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae;
                                                                                                                                                                                                                                                                       19 KNKVAQLKRKVR-----SLKDKAAELKQEVSRLENEIEDLKAKIGDLNNTSG 65
                                                                                                                                                                                                                                                                                                   ..
8
                                                                                                                                                                   21.2%; Score 80; DB 1; Length 102; 40.0%; Pred. No. 1.1; tive 8; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; L00602; AAA16595.1; -.
EMBL; D50617; BAA09230.1; -.
PIR; A44464; A49464.
PIR; S41804.
SGD; S0001886; SMC1.
InterPro; IPR003405; SMC_C.
InterPro; IPR03395; SMC_N.
Pfam; PF02463; SMC_C.
Pfam; PF02463; SMC_C: 1.
Mitosis; ATP-binding; Colled coll; Nuclear protein.
                                                                                                                          SEQUENCE 102 AA; 11555 MW; 1C9A45826A0B2AE3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Rel. 27, Last sequence update)
(Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 1225 AA
TIGR; MJ0334; -.
Interpro: IPR0061; REM_repeat.
Pfam; PP02185; HR1; 1.
Hypothetical protein; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
STRAIN-S288c / AB972;
MEDLINE-95400292; PubMed-7670463;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (Rel. 27, Created)
                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                   83
                                                                                                                                                                                                  Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-JUN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-0CT-1993
01-0CT-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                        SMC1_YEAST
P32908;
```

ij

7; Gaps

24; Indels

```
| | : |: :||||: | ::|| |: || |: || 367 GFRHQNSEGTGQAGDLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLEK 426
                                         3 GSHHHHHHGSMASGDLKNKVA-----QLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
  28.18; Prec ...
+1ve 15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M24720; AAA43111.1; ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE-89204899; PubMed-2705299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSSP, P03437; IHTM.
InterPro, IPRO01364; Hemagglutn.
Fram, PF00509; Hemagglutinin; 1.
PRINTS; PR00329; HEMAGGLUTNI2.
ProDom; PD000225; Hemagglutn; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              63580 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28.18;
                 18; Conservative
                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16
343
23
23
37
53
68
78
180
498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; C34064; HMIVE3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          180
300
498
565 AA;
     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Envelope protein;
                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11418;
                                                                                                                427 YVED 430
                                                                                       56 KIGD 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18;
                                                                                                                                                                          HEMA_IAHTO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNAL
                                                                                                                                                  RESULT 15
HEMA_IAHTO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                         à
                                                                 g
                                                                                        à
                                                                                                               셤
                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                        ä
                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE-89204899; PubMed=2705299;
MEDVINE-89204899; PubMed=2705299;
MEWOLUTION of the hemagglutinin of equine H3 influenza viruses.";
Virology 169:283-292(1989).
- I- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.
- I- SUBBUNT: HOMOTRAIMER, EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
- I- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
40 ATP (POTENTIAL).

89 COLLED COLL (POTENTIAL).

61 COLLED COLL (POTENTIAL).

61 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

64 ALA/ASP-RICH (DA-BOX).

73 S-L: IN TS MUTANT SMC1-2.

58 N->D: IN TS MUTANT SMC1-1.

141279 MW; B504017AA0ECCA8C CRC64;
                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N'LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                         12 SMASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIE----DLKAKIGDLNNT 63
                                                                                                            ; Score 80; DB 1; Length 1225; Pred. No. 14; 12; Mismatches 21; Indels
                                                                                                                                                                                                                                        HEMA_IAHAL STANDARD; PRT; 565 AA.
P16994; Q83991; Q83990;
01-AUG-1990 (Rel. 15, Created)
01-FEB-1991 (Rel. 17, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hemagglutinin Precursor [Contains: Hemagglutinin HAI chain;
                                                                                                                                                                                                                                                                                                                                        Influenza A virus (strain A/Equine/Algiers/72).
Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
Influenza A viruses; Influenzavirus A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001364; Hemagglutn.
Pfan; PF00509; Hemagglutinin; 1.
PRINTS; PR00329; HEMAGGLUTNI2.
PRODOM; PD000225; Hemagglutn; 1.
Envelope protein; Hemagglutinin; Glycoprotein; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BA533050DC3F186B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HEMAGGLUTININ HAI
HEMAGGLUTININ HAZ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M24721; AAA43100.1; ALT_SEQ.
PIR; D34064; HMIVE4.
HSSP; P03437; 1HTM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63831 MW;
                                                                                                            21.2%;
                                                                                                                         33.98;
                                                                                                                                    19; Conservative
                          1063
1061
1164
173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       565
23
37
53
78
180
  33 46
173 46
679 106
1057 106
1137 116
173 17
458 45
                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78
180
300
498
565 AA;
                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11393;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
CARBOHYD
CARBOHYD
                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
  NP_BIND
DOMAIN
                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SECUENCE
                                                             MUTAGEN
                         DOMAIN
DOMAIN
DOMAIN
                                                                           MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                      Matches
                                                                                                                                                                                                                       RESULT
                                                                                                                          Best
  FT FT FT SO
                                                                                                                                                            å
                                                                                                                                                                                 g
```

Length 565;

DB 1;

Score 79.5;

21,18;

Query Match

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Rawaoka Y., Bean W.J., Webster R.G.;
"Evolution of the hemagglutinin of equine H3 influenza viruses.";
"Influenza 169:283-292(1989).";
-i- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 GSHHHHHHGSMASGDLKNKVA-----QLKRKVRSLKDKAAELKQEVSRLENEIEDLKA 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC. . ) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          . . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21.1%; Score 79.5; DB 1; Length 565;
P17000; Q84002; Q84003; 01-AUG-1990 (Rel. 15, Created) 01-REP-1991 (Rel. 17, Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update) Hemagglutinin Precursor [Contains: Hemagglutinin HA2 chain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
Influenza A viruses; Influenzavirus A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15; Mismatches 24; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC...) (POT N-LINKED (GLCNAC..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hemagglutinin; Glycoprotein; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HEMAGGLUTININ HA1
                                                                                                                                                                                                                                                                                                                                                                                                                             Influenza A virus (strain A/Equine/Tokyo/71)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

367 GFRHQNSEGTGQAGDLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLEK 426

56 KIGD 59 : 1 427 YVED 430 οy

g

QQ

Search completed: December 18, 2002, 16:18:25 Job time: 3.13584 secs

-.

This Page Blank (uspto)

Q8wsp8 polyandroca Q8tfd6 emericella Q8x113 clostridium P91349 caenorhabd1

Q9v217 pyrococcus Q9y5q3 homo sapien 29lax3 streptococc

0980c6 Sulfolobus 084030 Influenza a 08404 influenza a 08404 influenza a 08134 halocynthia 09088 gallus gall 09088 gallus gall 09088 garlus gall 08xzz clostridium 08xzz clostridium 08xzz clostridium 08xzz clostridium 09xzz clostridium 09137 clostridium 09131 archaeoglob 09603 archaeoglob 09603 Influenza a 09003 drosophila 08007 pyrococcus 09752 schiztosach 09752 schistosama 090598 schistosama

OM protein

Run on:

Sequence:

Title:

Minimum DB Maximum DB

Database

Result

è

Searched:

```
MEDIATRE-201703;
MEDIATRE-20448953; PubMed-10992499;
Hellingshead S.K., Becker R., Briles D.E.;
"Diversity of PepA; mosalc genes and evidence for past recombination in Streptococcus pneumonlae.";
Infect. Immun. 68:5889-5900(2000).
Infect. Immun. 68:5889-5900(2000).
Interpro; IPR002965; P_rich_extensn.
PRINTS; PR0121; PRHIEXTENSN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Streptococcus pneumoniae.
acteria: Firmicutes: Bacillus/Clostridium group; Lactobacillales; Streptococcaccaes: Streptococcaccaes: Streptococcaccae: Streptoc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ouery Match 24.1%; Score 91; DB 2; Length 481; Best Local Similarity 37.2%; Pred. No. 4.7; Matches 16; Conservative 15; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER 481 481
SEQUENCE 481 AA; 53500 MW; EA3C66445EFECE2B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17 DLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGD 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 481 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
091AX3
09Y217
09Y217
09Y217
09Y217
08TL13
P91349
09Y010
08OT94
09W21
09W21
09W21
09W21
09C79
09C79
09C79
09C79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9U7E3
O54615
Q9HQM8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09Y7S2
09TY57
096398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                            7
10
10
10
10
                                                                                                                                                                                                                                                                                                                                                                                                16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9LAX5;
01-OCT-2000 (TrEMBLrel. 15,
01-OCT-2000 (TrEMBLrel. 15,
01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                451
1188
212
212
240
240
2244
396
802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PspA (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                    81.5
81.5
81.5
81.8
81
81
81
81
81
81
80.5
                                                                                                                                                                                                                                                         RESULT 2
Q8TV66
ID Q8TV66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9LAX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 1
Q9LAX5
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
7444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
7444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
7444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
74444
744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Oftv66 methanopyru

OgBuk5 brachydanio

O73679 brachydanio

O980k3 brachydanio

O91891 gallus gall

O91891 gallus gall

O57342 coturnix co

O66901 aquifex aeo

O96691 brachydanio

O31405 gallus gall

O81x14 methanopyru

O82290 gallus gall

O91111 homo saplen

O31620 gallus gall

O91611 homo saplen

O31620 gallus gall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9lax5 streptococc
                                                                                                                                                                                (without alignments)
3509.535 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                        December 18, 2002, 16:13:04; Search time 4.34459 Seconds.
                                                                                                                                                                                                                                                                                               74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Description
                                                                                                                                                                                                                                                                                               1 MRGSHHHHHHGSMASGDLKN.....AKIGDLNNTSGIRRPAAKLN
                   GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                 671580 seqs, 206047115 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                            protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8TV66
Q98UK5
Q98UK3
Q91891
Q57342
O66901
Q986B1
Q31405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08TXI4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9H1F1
Q31620
Q9LAX2
                                                                                                                                                                                                                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9LAX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5: Sp_invertebrate:*
6: Sp_manmal:*
7: Sp_mhc:*
8: Sp_organelle:*
9: Sp_phage:*
10: Sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2
113
113
113
114
117
117
                                                                                                                                                                                                                                                    US-09-490-291-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SPTREMBL_21:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4: sp_human:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
Match I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24.1
223.7
233.7
223.3
222.8
222.8
222.7
222.7
222.5
222.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Perfect score:
                                                                                                                                                                                                                                                                                                                                             Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score
                                                                                                               •
```

ö

Gaps

ö

```
275 RRILKNRGYAQSCRFKRVQQKHLLENEKTQLINQVEQLKQEINRLARERDAYKLKCEKLT 334
                                                   335 GANGER 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           335 GANGER 340
                          62 NTSGIR 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62 NTSGIR 67
                                                                                                              073679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      098UK3
                                                                                        RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
                                                                                                    073679
                          ð
                                                   g
                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3;
                                                                                      Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Brachydanio rerio (Zebrafish) (Zebra danio).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 HHHHHH-----ASGDLKNKVAQLKRK 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----AAELKQEVSRLENEIEDLKAKIGDLN 61
                                                                                                                                             STRAIN-AV19 / DSM 6324 / JCM 9639;
MEDLINE-21927647; PubMed=11930014;
Slesarev A.L., Mazhevaya K.V., Makarova K.S., Polushin N.N.,
Slcherbinian O.V., Shakhova V.V., Belova G.I., Aravind L.,
Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
Malykh A.G., Koonin E.V., Kozyavkin S.A.;
The complete genome of hyperthormophile Methanopyrus kandleri AV19
Proc. Natl. Acad. Sci. US.A. 99:4644-4649(2002).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "Isolation, Characterization, and Expression Analysis of Zebrafish Large Mafs.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-21064923; PubMed-11134968;
Kajihara M., Kawauchi S., Kobayashi M., Ogino H., Takahashi S.,
                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5; DB 13; Length 356;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 91; DB 17; Length 615; Pred. No. 6;
                                                                                                                                                                                                                                                                                                                                                                            15 SGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNNT 63
                                                                                                                                                                                                                                                                                                                                    9; Mismatches 21; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                     615 AA; 66941 MW; 77729702806BC922 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; ABOO6322; BAB21102.1; -.
InterPro; IPR004825; TE_bZIP.
InterPro; IPR004825; TE_bZIP.
InterPro; IPR00131; DZIP_MAf; 1.
SMART; SM00338; BRLZ; 1.
SEQUENCE 356 AA; 40233 MW; DE4C96B62C058865 CRC64;
          01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
Predicted secreted protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         0980K5;
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
Transcription factor MafB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   356 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23.7%; Score 89.5; D
22.2%; Pred. No. 4.7;
.ive 13; Mismatches
                                                                                                                                                                                                                                                                                                             24.1%;
38.8%;
                                                                                                                                                                                                                                                                                                                                      19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                       Methanopyrus kandleri.
                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                       Complete proteome.
SEQUENCE 615 AA;
                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                             NCBI_TaxID-2320;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29 VRSLKDK---
                                                                                                Methanopyrus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Yasuda K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                0980K5
                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
ð
                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
```

```
[1]
SEQUENCE FROM N.A.
MEDLINE-98165393; PubMed-9425134;
Moens C.B., Cordes S.P., Giorgianni M.W., Barsh G.S., Kimmel C.B.;
"Equivalence in the genetic control of hindbrain segmentation in fish and mouse.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              275 RRILKNRGYAQSCRFKRVQQKHLLENEKTQLINQVEQLKQEINRLARERDAYKLKCEKLT 334
                                                                                                          VAL OR VALENTINO.

Brachydanio rerio (Zebrafish) (Zebra danio).

Brachydanio rerio (Zebrafish) (Zebra danio).

Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

Cyprinidae; Danio.

NCBI_TaxID=7955;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 HHHHHH-----ASGDLKNKVAQLKRK 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------AAELKQEVSRLENEIEDLKAKIGDLN 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Brachydanio rerio (Zebrafish) (Zebra danio).
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Actinopterygil; Neopterygil; Teleostel; Ostariophysl; Cypriniformes;
Cyprinidae; Danio.
NCBI_TaxID=7955;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Yasuda K.; "Isolation, characterization, and Expression Analysis of Zobrafish
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [1]
SEQUENCE FROM N.A.
MEDLINE-21064923; PubMed-11134968;
Kajihara M., Kawauchi S., Kobayashi M., Ogino H., Takahashi S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 356;
                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR004827; TF_bZIP.
InterPro; IPR004826; TF_Maf.
InterPro; IPR0131; bZIP_Maf; 1.
SMART; SM00338; BRLZ; 1.
SEQUENCE 356 AA; 40243 MW; 07420DB0F6CD08F1 CRC64;
                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 13;
356 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23.7%; Score 89.5; D 22.2%; Pred. No. 4.7; ive 13; Mismatches
                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                               Development 125:381-391(1998).
EMBL; AF006641; AAC18821.1; -.
ZFIN; ZDB-GENE-980526-515; val.
                                 01-AUG-1998 (TrEMBLrel. 07, 01-AUG-1998 (TrEMBLrel. 07, 01-MAR-2002 (TrEMBLrel. 20,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JUN-2001 (TrEMBLrel. 17, 01-JUN-2001 (TrEMBLrel. 17, 01-MAR-2002 (TrEMBLrel. 20,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29 VRSLKDK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 23.7%
Best Local Similarity 22.2%
Matches 28; Conservative
                                                                                                    Transcription factor Val.
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Large Mafs.";
```

÷,

```
Coturnix coturnix japonica (Japanese quail).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                        28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Complete proteome.
SEQUENCE 259 AA;
                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                            NCBI_TaxID=93934;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               HYPB OR AQ_671.
Aquifex aeolicus.
                                                                                                                                                                                                             SECUENCE
                                                                                                                                                                                                                                                                                                                                                                                                 066901
                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
O66901
                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
    g
                                                                                                                                                                                                                                                                                                                       οχ
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                              ð
                                                                                                              ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .,
                                                                                                                                                                                                                                                                                      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
bZIP transcription factor L-MAR (Fragment).
Gallus gallus (Chicken).
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                               Gaps
                                                                                                                                   4 SHHHHHHG------SMASGDL-----35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      :|||||||
39 AHHHHHHHHLRLEERFSDDQLVSMSVRELNRQLRGFSKEEVIRLKQKRRTLKNRGYAQSC
                                                                                                                                                                                                                                                                                                                                                                                                           Zhou H., Lamont S.J.; "Genetic variation among chicken lines and mammalian species in specific genes.";
                                                                                                              38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23.5%; Score 88.5; DB 13; Length 146; 27.5%; Pred. No. 2.3; ive 11; Mismatches 14; Indels 49
                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
STRAIN-BREED FAYOUMI, AND BREED G-B1 LEGHORN; TISSUE-SPLEEN;
                                                                                        Length 315;
                                                                                                            20; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases. EMBL, AF221558; AAF71240.1; -. EMBL, AF221557; AAF71239.1; -. Interpro; IPR004827; TF_AZIP. Interpro; IPR004826; TF_Maf. Pfan; PF03131; bZIP_Maf; 1. SM00338; BRLZ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          146 AA; 17579 MW; 382BAD1D098E1DCC CRC64;
                                                                   E0147093BAC67CE2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99 RYKRVQQRHILENEKCQLQSQVEQLKQEVSRLAKERDLYKEK 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-JUN-1998 (TrEMBLrel. 20, Last annotation update)
bzIP transcription factor MafA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              286 AA
                                                                                                                                                                                          |: | |::::| || : |
SCRYKRVQQRHMLESEKCTLQSQVEQLKQDVARL 272
                                                                                                            12; Mismatches
                                                                                                                                                                              ------LKQEVSRLENEIEDLKAKIGDL
                                                                                        Score 89;
Pred. No.
                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
J. Biochem. 129:139-146(2001).
EMBL; AB006324; BAB21104.2; -.
InterPro; IPR004825; TF_Maf.
Pfam; PF03131; DZIP_Maf; 1.
SMART; SM00338; BRLZ; 1.
SEQUENCE 315 AA; 35705 MW; E
                                                                                        23.6%; 25.5%;
                                                                                                                                                                                                                                                                 Q91891;
01-OCT-2000 (TrEMBLrel. 15,
                                                                                 Query Match
Best Local Similarity 25.59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 27.5 nes 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 146
                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=9031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            057342
                                                                                                                                                                                                                                                        091891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loc
Matches
                                                                                                                                                                            39
                                                                                                                                                                                                 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 7
057342
ID 0573
AC 0573
AC 01-J
DT 01-J
DT 01-M
DE bZIP
GN MAFP
                                                                                                                                                                                                                                  RESULT 6
091891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
    ð
                                                                                                                                                       qq
                                                                                                                                                                              ò
                                                                                                                                                                                                 셤
```

```
ج
..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ë
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE=98337467; PubMed=9674710;
Benkhelifa S., Provot S., Lecoq O., Pouponnot C., Calothy G.,
Benkhelifa S., Provot S., Lecoq O., Pouponnot C., Calothy G.,
"mafA, a novel member of the maf proto-oncogene family, displays
developmental regulation and mitogenic capacity in avian neuroretina
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-VF5;
MEDLINE-98196666; PubMed-9537320;
MEDLINE-98196666; PubMed-9537320;
Deckett G., Warren P.V., Gassterland T., Young W.G., Lenox A.L.,
Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
Fridman R.A., Short J.M., Olson G.J., Swanson R.V.;
The complete genome of the hyperthermophilic bacterium Aquifex
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SHHHHHHG------SMASGDL------35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -------EVS 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        49;
                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 16; Length 259;
                                                                                                                                                                                                                                                                                                                                                                                                                                           23.5%; Score 88.5; DB 13; Length 2
27.5%; Pred. No. 4.5;
ative 11; Mismatches 14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
NCBI_TaxID=63363;
                                                                                                                                                                                                                                                                                                                                                                                    286 AA; 32463 MW; 0D4ED7BF08D91397 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29015 MW; 4522AC8CFA53806C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             217 RYKRVQQRHILENEKCQLQSQVEQLKQEVSRLAKERDLYKEK 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AuG-1998 (TrEMBLrel. 07, Created)
01-AuG-1998 (TrEMBLrel. 07, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
Hydrogenase expression/formation protein B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mature 392.353-358(1998).

EMBL; AE000701; AAC05860.1; -
InterPro; IPR004392; HypB.
InterPro; IPR002894; HypB_UreG.
InterPro; IPR00822; Znf_C2H2.
Pfam; PF01495; HypB_UreG.
InterPro; IPR00073; hypB.InterPro; IPR00073; hypB; I.
PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----ENEIEDLKAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 HHHHHHGSMASGDLKNKVAQLKRKVRSLKDKAAELKQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               259 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 23.3%; Score 88; DB Best Local Similarity 28.6%; Pred. No. 4.5; Matches 26; Conservative 12; Mismatches
                                                                                                                                                                                                      Oncogene 17:247-254(1998).
EMBL; AF034693; AAC60377.1; -
InterPro; IPR004827; TF_bZIP.
InterPro; IPR004826; TF_Maf.
Pfam; PF03131; bZIP_Maf; I.
SMART; SM00338; BRLZ; I.
```

3

Length 155;

DB 7;

à 요 

```
1 1 1 1 1 155 AA; 18201 MW; 7FC66EF2584F3978 CRC64;
                                              22.8%; Score 86; 27.5%; Pred. No.
                                                                                                                                                                                                                                                                                                                    01-JUN-2002 (TrEMBLrel, 21, 01-JUN-2002 (TrEMBLrel, 21, 01-JUN-2002 (TrEMBLrel, 21, Uncharacterized protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-2002 (TrEMBLrel. 21, 01-JUN-2002 (TrEMBLrel. 21, 01-JUN-2002 (TrEMBLrel. 21,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMC1-family ATPase involved
                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                    Methanopyrus kandleri.
                                            Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 20; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Complete proteome. SEQUENCE 609 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                  Methanopyrus.
NCBI_TaxID=2320;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-2320;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SECC OR MK0690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Methanopyrus
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                          DBTXA4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QBTXI4;
                                                                                                                                                                                                                                                                                        Q8TXA4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OSTXI4
                                                                                                                                                                                                                                                         RESULT 11
OBTXA4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q8TXI
FT
                                                                                                                                           g
                                                                                                                                                                            à
                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4
                                                                                                                                                                                                                                                                                                                        MEDLINE-99173794; PubMed-10072793; Schwarzstein M., Kirn A., Haffter P., Cordes S.P.; Schwarzstein M., Kirn A., Haffter P., Cordes S.P.; Expression of zkml2, a homologue of the Kimll/val segmentation gene, during embryonic patterning of the zebrafish."; Mech. Dev. 80:223-226(1999). EMBL; AP109781; AAD26(1111; -. ZFIN; ZDB-GENE-010605-2; maf. InterPro; IPR004826; TF_bZIP. InterPro; IPR004826; TF_Maf. InterPro; IPR004826; TF_Maf. SWART; SM00318; BRLZ; 1. SWART; SM00318; BRLZ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 176 HHHHHHHHHGHQOHQTPASRPSSTNSLSPITRQPPDDRFSDDQLVTMSVRELNRHVRGVSKE 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Metazoa; Chórdata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                Brachydanio rerio (Zebrafish) (Zebra danio).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygil; Neopterygil; Teleostel; Ostariophysi; Cypriniformes;
Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21 KVAQLKRKVRSLKDKA------51 KVAQLKRKVRSLKDKA------55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----KN 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [1] SEQUENCE FROM N.A. MEQUENCE FROM N.A. MEDIJINE-90077532; PubMed-2592020; Kaufman J., Salomonsen J., Skjoedt K.; Raufman J., Salomonsen H., Skjoedt K.; B-G CDNA clones have multiple small repeats and hybridize to both
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23.1%; Score 87; DB 13; Length 324; 24.2%; Pred. No. 6.9; ive 11; Mismatches 19; Indels :
                                                                                                                                                  01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-NAR-2002 (TrEMBLrel. 20, Last annotation update)
Basic domain leucine zipper transcription factor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                         324 AA.
                          RLENEIEDLK--AKIGDLNNTSGIRRPAAKL 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   chicken MHC regions.";
Immunogenetics 30:440-451(1989).
EMBL; M27668; AAA48624.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q31405
Q31405;
01-NOV-1996 (TEMBLEEL. 01, 0
01-NOV-1996 (TEMBLEEL. 01, 0)
01-DEC-2001 (TEMBLEEL. 19, 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MHC B-G antigen (Fragment). Gallus gallus (Chicken).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 ННННННВ-----
                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    56 KIGDLNNTSGIR 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=9031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                       09W6B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gallus
         45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 10
                                                                                      RESULT
09W6B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             931405
```

合 ò

ò g ó g 

```
Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Methanopyrus kandleri.
Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
        Gaps
                                                     7 HHHHGSMASGDLKNKVA------QLKRKVRSLKDKAAELKQEVSRLENEIED 52
                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N., Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L., Natalab D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O., Malykh A.G., Koonin E.V., Kozyavkin S.A., "The complete genome of hyperthermophile Methanopyrus kandleri AV19 and monophyly of archaeal methanogens."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15 SGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDLNNTSGIRRPAAKL 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
STRAIN=AV19 / DSM 6324 / JCM 9639;
SHDLINE-21927647; PubMed=11930014;
SIESATEV A.I., Mezhevas K.V., Masarova K.S., Polushin N.N., Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L., Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
           18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22.8%; Score 86; DB 17; Length 609; 33.9%; Pred. No. 16; tive 15; Mismatches 22; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    609 AA; 69552 MW; AB10C9780DC5AD78 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
in DNA repair.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
EMBL; AE010369; AAM01985.1; -.
                                                                                                                                                                                                                                                                                                                                                              609 AA.
17; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
STRAIN-AV19 / DSM 6324 / JCM 9639;
MEDLINE-21927647; Pubmed-11930014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                 53 LKAKIGDLNNTSGIRRPAAK 72
                                                                                                                                                                                                        79 VEKHL----KKTGRRAPNLK 94
```

ä

S

```
Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
                                       EMBL; AL035665; CAB75863.1; ...
InterPro; IPR004827; TF_bZIP.
InterPro; IPR04826; TF_maf.
Pfam; PF03131; bZIP_maf; 1.
SMART; SM00338; BRLZ; 1.
    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                          RESULT 15
    ď
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                       à
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3;
                                                                                                                                   ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                157 AHHHHHHHHLRLEERFSDDQLVSMSVRELNRQLRGFSKEEVIRLKQNRRTLKNRGYAQSC 216
                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 SHHHHHHG------SMASGDL-----35
Malykh A.G., Koonin E.V., Kozyavkin S.A.;
"The complete genome of hyperthermophile Methanopyrus kandleri AV19
and monophyly of archaeal methanogens.";
Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
EMBL; AE010362; AAM01904.1; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens (Human),
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
NCBL_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   217 RYKRVQQRHILENEKCQLQSQVEQLKQEVSRLAKERDLYKEKYEKLAARGFPREPS 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 36 ------36 DINNTSGIRRPA 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-WAR-2001 (TrEMBLrel. 16, Created)
01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DJ644Ll.1 (Kreisler (mouse) maf-related leucine zipper homolog).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 22.7%; Score 85.5; DB 13; Length 286; Best Local Similarity 25.9%; Pred. No. 8.1; Matches 30; Conservative 12; Mismatches 25; Indels 49
                                                                                                       22.8%; Score 86; DB 17; Length 876; 45.5%; Pred. No. 23; tive 8; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ogino H., Yasuda K.;
"Induction of lens differentiation by activation of a bZIP
"Induction factor, L-Maf.";
Science 280:115-118(1998).
EMBL, AF034570; AAC15781.1; -.
InterPro; IPR004827; TF_bZIP.
InterPro; IPR004826; TF_Maf.
Pfam; PF03131; bZIP_Maf.
SMART; SM00338; BRLZ; 1.
                                                                              876 AA; 103706 MW; D651E682271FD383 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 286 AA; 32445 MW; 15DBD7B9BBDFA522 CRC64;
                                                                                                                                                                          193 DLKGSKKELKRVERELE----ELKREVKELEPEVEELKERLNEL 232
                                                                                                                                                           17 DLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDL 60
                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           323 AA
                                                                                                                                                                                                                                                                             (TrEMBLrel, 05, Created) (TrEMBLrel, 05, Last seq
                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE-LENS;
MEDLINE-98192815; PubMed-9525857;
                                                                                                                                                                                                                                                                                                       01-MAR-2002 (TrEMBLrel. 20, Last
bZIP transcription factor L-Maf.
                                                                                                                  Local Similarity 45.5 res 20; Conservative
                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                              Gallus gallus (Chicken)
                                                                 Complete proteome. SEQUENCE 876 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9031;
                                                                                                                                                                                                                                                               042290;
01-JAN-1998
                                                                                                                                                                                                                                                                                         01-JAN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9H1F1;
                                                                                                                                                                                                                                                  042290
                                                                                                                                                                                                                                                                                                                                                                                      Gallus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09H1F1
                                                                                                                                                                                                                          RESULT 13
042290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
                                                                                                                                                                                   g
                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
```

```
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-91239571; PubMed=1903541;
Miller M.M., Goto R., Young S., Chirivella J., Hawke D., Miyada C.G.;
"Immunoglobulin variable-region-like domains of diverse sequence
"Immunoglobulin variable-region-like domains of the chicken.";
Proc. Natl. Acad. Sci. U.S.A. 88:4377-4381(1991).
EMBL: MG1862; AAA48627.1; -.
Interpro; IPR003600; Ig_like.
Interpro; IPR003600; Ig_like.
                                                                                                                                                                                                                                Gallus gallus (Chicken).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                162 HHHHHQASPPPSSAASPAQQLPTSHPGPGPHATASATAAGGNGSVEDRFSDDQLVSMSV 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7 HHHHGSMASGDLK------NKVAQLKRKVRSLKDKAAELKQEVSRLENEIED 52
                                                                                                                                                                                                        16 -----GDLKNKVAQLKRKVRSLKDK------AAELK
                                                                                   85;
                                        DB 4; Length 323;
.2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22.5%; Score 85; DB 7; Length 321; llarity 26.8%; pred. No. 10; Conservative 18; Mismatches 24; Indels
                                      Score 85.5; DB 4; Length 32
Pred. No. 9.2;
9; Mismatches 19; Indels
323 AA; 35792 MW; A0F3C09F8936CB16 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      321 AA; 36360 MW; 7F25BD29E39860C7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                    321 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: December 18, 2002, 16:21:03 Job time : 6.34459 secs
                                                                                                                                                                                                                                                                                                                                282 QEVSRLARERDAYKVKCEKLAN-SGFR 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                       41 QEVSRLENEIEDLKAKIGDLNNTSGIR 67
                                                                                                                         5 нннннн------СЅМАЅ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        245 VEKHLKKI----GIRAPNLKLH 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               53 LKAKIGDLNNTSGIRRPAAKLN 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (TrEMBLrel. 01, C
(TrEMBLrel. 01, I
(TrEMBLrel. 19, I
                                      22.78;
23.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity 23.18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         B-G protein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
es 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=9031;
                                                                                                                                                                                                                                                                                                                                                                                                                                031620
Q31620;
01-NOV-1996
01-NOV-1996
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
SEQUENCE
```

THIS PAGE BLANK (USPTO)